CHARTER MARK

The Patent Office

PCT/GB99/02406

09/744502

INVESTOR IN PEOPLE

GB99

02406

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

E5KU

REC'D 26 AUG 1999

WIPO PCT

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed Andrew Gersey

Dated 13 August 1999

For official use

9816265.4

24 JUL 1998

24 JUL 1998
LONDON

27JUL98 E378661-7 D03312
P01/7700 25.00 - 9816265.4

Your reference 230P79179

Notes

Please type, or write in dark ink using CAPITAL letters. A prescribed fee is payable for a request for grant of a patent. For details, please contact the Patent Office (telephone 071–438 4700).

Rule 16 of the Patents Rules 1990 is the main rule governing the completion and filing of this form.

❷ *Do not give trading styles, for example, 'Trading as XYZ company', nationality or former names, for example, 'formerly (known as) ABC Ltd' as these are not required.*

Warning

After an application for a Patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977 and will inform the applicant if such prohibition or restriction is necessary. Applicants resident in the United Kingdom are also reminded that under Section 23, applications may not be filed abroad without written permission unless an application has been filed not less than 6 weeks previously in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction revoked.

The Patent Office

Request for grant of a Patent

Form 1/77

Patents Act 1977

❶ Title of invention

1 Please give the title of the invention: Protection of plants

❷ Applicant's details

☐ **First or only applicant**

2a If you are applying as a corporate body please give:

Corporate name: Instituto Tecnico Agronomico Provincial, S.A.

Country *(and State of incorporation, if appropriate)*: SPAIN

2b If you are applying as an individual or one of a partnership please give in full:

Surname

Forenames

2c **In all cases**, please give the following details:

Address: Albacete

SPAIN

UK postcode *(if applicable)*

Country

ADP number *(if known)*: 07483737001

2d, 2e and 2f: If there are further applicants please provide details on a separate sheet of paper.

Second applicant *(if any)*

2d If you are applying as a corporate body please give:

Corporate name

Country *(and State of incorporation, if appropriate)*

2e If you are applying as an individual or one of a partnership please give in full:

Surname

Forenames

2f **In all cases**, please give the following details:

Address

UK postcode *(if applicable)*

Country

ADP number *(if known)*

❸ An address for service in the United Kingdom must be supplied

Please mark correct box

❸ Address for service details

3a Have you appointed an agent to deal with your application?

Yes [X] No [ ] ➡ ***go to 3b***

please give details below

Agent's name: Marks & Clerk

Agent's address: 57-60 Lincoln's Inn Fields
LONDON
WC2A 3LS.

Postcode

Agent's ADP number: 18001

3b: If you have appointed an agent, all correspondence concerning your application will be sent to the agent's United Kingdom address.

3b If you have not appointed an agent please give a name and address in the United Kingdom to which all correspondence will be sent:

Name

Address

Postcode

Daytime telephone number *(if available)*

ADP number *(if known)*

❹ Reference number

4 Agent's or applicant's reference number *(if applicable)*

230P79179

❺ Claiming an earlier application date

Please mark correct box

5 Are you claiming that this application be treated as having been filed on the date of filing of an earlier application?

Yes ☐ No ☒ ➡ ***go to 6***

⬇ ***please give details below***

❑ number of earlier application or patent number

❑ filing date

(day month year)

❑ and the Section of the Patents Act 1977 under which you are claiming:

Please mark correct box

15(4) (Divisional) ☐ 8(3) ☐ 12(6) ☐ 37(4) ☐

❻ Declaration of priority

❻ If you are declaring priority from a PCT Application please enter 'PCT' as the country and enter the country code (for example, GB) as part of the application number.

Please give the date in all number format, for example, 31/05/90 for 31 May 1990.

6 If you are declaring priority from previous application(s), please give:

Country of filing	Priority application number *(if known)*	Filing date *(day, month, year)*

❼ The answer must be 'No' if:
● any applicant is not an inventor
● there is an inventor who is not an applicant, **or**
● any applicant is a corporate body.

❼ Inventorship

7 Are you (the applicant or applicants) the sole inventor or the joint inventors?

Please mark correct box

Yes [ ] No [X] ➡ ***A Statement of Inventorship on Patents Form 7/77 will need to be filed (see Rule 15).***

❽ Please supply duplicates of claim(s), abstract, description and drawing(s).

❽ Checklist

8a Please fill in the number of sheets for each of the following types of document contained in this application.

Continuation sheets for this Patents Form 1/77 —

Claim(s)	—	Description	26
Abstract	—	Drawing(s)	4 + 4 (FORMAL)

8b Which of the following documents also accompanies the application?

Please mark correct box(es)

Priority documents *(please state how many)*	—
Translation(s) of Priority documents *(please state how many)*	—
Patents Form 7/77 – Statement of Inventorship and Right to Grant *(please state how many)*	—
Patents Form 9/77 – Preliminary Examination/Search	—
Patents Form 10/77 – Request for Substantive Examination	—

…our appointed agent (see Rule 90 of the Patents Rules 1990) must sign this request.

❾ Request

I/We request the grant of a patent on the basis of this application.

Please sign here ➡

Signed Marks & Clerk Date 24 / 7 / 98 *(day month year)*

A completed fee sheet should preferably accompany the fee.

Please return the completed form, attachments and duplicates where requested, together with the prescribed fee to:

❑ ***The Comptroller***
The Patent Office
Cardiff Road
Newport
Gwent
NP9 1RH

or

❑ ***The Comptroller***
The Patent Office
25 Southampton Buildings
London
WC2A 1AY

Document #: 288426D

Protection of Plants

OBJECT OF THE INVENTION

The present invention describes a new species of binucleated *Rhizoctonia* suited for use in plant protection.

SUMMARY OF THE INVENTION

The present invention describes a new species of binucleated *Rhizoctonia* able to protect different kind of plants, like tomatoes, onions, grass, pine trees, and others, from infections produced by different pathogenic fungi among them *Fusarium*, *Verticillum*, *Phythophtora* and pathogenic *Rhizoctonia*, as well as yeast and pathogenic bacteria and the appropriated methodology for the molecular identification of this new species, its detection based on probes specially designed and its practical application for the protection of crops.

Thus the present invention uses the ITS data to distinguish isolates and in particular to identify pathogenic and candidate protective isolates belonging to the binucleated group.

The invention also provides specific PCR primers for the ITS molecular analysis of pathogenic and protective strains of *Rhizoctonia*.

PREFERRED EMBODIMENTS

The protective species can be utilised in a form of powder obtained after fermentation of the micro-organisms with subsequent drying of the filtered broth , applied to the seed before seeding or applied to the soil before seeding or both , and will preclude the growth of the pathogen *Rhizoctonia* protecting the plant from the infection.

In the present invention, the ITS sequence is employed as a marker to distinguish strains and correspondingly classify the pathogenicity of isolates of *Rhizoctonia*. The ITS information enables phylogenetic location of protective isolates, and design of specific primers for the identification of new isolates.

SUMMARY OF FIGURES

Figure 1 is a schematic representation of the Internal Transcribed Spacer (ITS) region of ribosomal DNA of *Rhizoctonia*.

Figure 2 is a phylogenetic tree for the results obtained in the Examples.

Figure 3 shows photographic evidence of the protective effect of RBN isolates B and F2 on tomato seedlings.

3.a.- Shows the plates corresponding to the tomato seedling as negative control and the plate with tomato seedling with the pathogenic strain Me 8-2.

3.b.- Shows the negative control with the protective isolates without pathogen.

3.c.- Protective effect of F2 on Me 8-2 action, both simultaneous and subsequently.

3.d.- Protective effect of B on Me 8-2 action , both simultaneous and subsequently.

EXAMPLES OF THE INVENTION

In this study we sequenced the ITS - 5.8S regions of 13 binucleate isolates of *Rhizoctonia* isolated from Albacete (South-East of Spain), 9 binucleate isolates from CBS (Central Bureau Schimmelcultures, Baarn, Delft, Netherlands) collection, 3 multinucleate isolates from Albacete and multinucleate isolates belonging to AG 2. The later multinucleate isolates were included to determine the sequence variation in rDNA regions and to assess this variation as a possible tool for identifying groups into *Rhizoctonia* spp. More generally, the isolates from collections were included for comparison purposes and completeness of the phylogenetic tree.

Isolation of *Rhizoctonia* spp. from the field.

Samples of saffron, alfalfa and pine seedlings were collected at different locations in Albacete. South East of Spain. Isolation of *Rhizoctonia* spp. from these samples was performed by washing saffron roots, alfalfa stems and roots and whole pine seedlings with tap water to remove adhering soils particles and plating plant segments on 2% water agar amended with chloramphenicol (250 mg/ml). The cultures were characterised as *Rhizoctonia* by observing hyphal and colony morphology and vegetative features microscopically, such as hyphae with dolipore septa, basally constricted branching near the distal septum of cells, and absence of clamp connections and conidia. Cultures were maintained and stored at room temperature on potato-dextrose-agar [PDA, Oxoid LTD, U.K.]. The number of nuclei per hyphal cells was determined using a staining method described by Julian *et al.* (Julian *et al.*, 1997). Briefly, 3 mm^2 agar plugs of inoculum were placed in petri dishes containing liquid complete medium [CM: 5 g malt extract (Merck, Darmstadt, Germany), 5 g yeast extract (Difco, Detroit, USA) and 5 g glucose anhydrous (Merck, Darmstadt, Germany) per litter] and incubated for 3 d at 230 C in the dark. The resulting mycelial colony was transferred to a microscope slide, excess medium was removed and a drop of 50 mg/ml Hoechst 33342 (Molecular Probes Europe, Leiden, The Netherlands) was added. After a 15 min incubation in the dark, the staining solution was removed and 50% glycerol was added. A coverslip was placed on sample and the number of nuclei was determined by fluorescence microscopy using an Olympus BHS microscope supplemented with a BH2-RFC reflected light fluorescence attachment (with dichroic mirrors BH2-DMU for UV excitation).

Other *Rhizoctonia* isolates

Binucleate *Rhizoctonia* spp. isolates have been obtained from the CBS collection and are listed in Table 2.

Table 1. Original isolates described and sequenced for the first time in the invention

Isolates	Nuclear condition	Identification	Isolated from	Geographical origin
F1	Binucleate	Unknown	Saffron	Albacete (Spain)

F2	Binucleate	Unknown	Saffron	Albacete (Spain)
F3	Binucleate	Unknown	Saffron	Albacete (Spain)
F5	Binucleate	Unknown	Saffron	Albacete (Spain)
F6	Binucleate	Unknown	Saffron	Albacete (Spain)
F7	Binucleate	Unknown	Saffron	Albacete (Spain)
S1	Binucleate	Unknown	Saffron	Albacete (Spain)
S2	Binucleate	Unknown	Saffron	Albacete (Spain)
S3	Binucleate	Unknown	Saffron	Albacete (Spain)
S4	Binucleate	Unknown	Saffron	Albacete (Spain)
S5	Binucleate	Unknown	Saffron	Albacete (Spain)
S6	Binucleate	Unknown	Pinus pinaster	Albacete (Spain)
S7	Multinucleate	Unknown	Pinus pinaster	Albacete (Spain)
A1	Multinucleate	Unknown	Alfalfa	Albacete (Spain)
A2	Multinucleate	Unknown	Alfalfa	Albacete (Spain)
B	Binucleate	Unknown	Alfalfa	Albacete (Spain)

Table 2 Reference isolates belonging to the Binucleated Group obtained from different collections

Isolates	Nuclear Condition	Identification	Isolated in	Geographical Origin
C1	Binucleate	Ceratobas. Cornigerum	Festuca sp.	Pennsylvania USA
C2	Binucleate	Cerarobas. Cornigerum	lead of Pittosporum	Pennsylvania USA

C4	Binucleate	Ceratobas. Cornigerum	Juniperus sp.	North Carolina USA
C5	Binucleate	Ceratobas. Cornigerum	Taxus sp.	Rhode Island USA
C6	Binucleate	Ceratobas. Cornigerum	Erigeron canadiensis	Pennsylvania USA
C8	Binucleate	Ceratobas. Cornugerum	Pittosporum sp.	Florida USA
C11	Binucleate	Ceratobas. Cerealis	Secale cereale	Wageningen. NL
C12	Binucleate	Ceratobas. Cerealis	Secale cereale	Germany
C13	Binucleate	Ceratobas. Cerealis	Triticum aestivum	Germany
17-329	Binucleate AG-A	Ceratobas. Corgigerum	Strawberry	Saga. Japan
17-326	Binucleate AG-A (CAG-2)	Ceratobas. Cornigerum		USDA
16-321	Binucleate AG-BO	Ceratobas. Cornigerum	Patata	Tokushima. Japan
17-328	Binucleate AG-B6			USDA
16-320	Binucleate AG-C	Ceratobas. Cornigerum	Gimnadenia conopsea	Hokkaido. Japan
16-319	Binucleate AG-D	Ceratobas. Gramineum	Wheat	Hokkaido. Japan
4-7 b	Binucleate AG-E	Ceratobasidium sp.	Flax	Hokkaido. Japan
17-331	Binucleate AG-G	Ceratobasidium sp.	Peanut	Chiba. Japan
16-315	Binucleate AG-I	Unknown Artemisia	Tokaydo. Japan	
16-314	Binucleate AG-J	Ceratobasidium sp.	Soil	Kanto. Japan
16-317	Binucleate AG-K	Ceratobasisium sp.	Onion	Hokkaido. Japan

.9-363	Binucleate AG-L	Ceratobasidium sp.	Soil	Fukuoka, Japan
19-368	Binucleate AG-N	Unknown	Soil	Tochigi, Japan
19-367	Binucleate AG-P	Ceratobas. Cornigerum	Tea	Shizuoka, Japan
19-365	Binucleate AG-R	Ceratobasidium sp.	Cucumis	Georgia, USA
19-366	Binucleate AG-S(CAG-7)	Ceratobas. sp.	Pinosporum	Florida, USA

Table 3 Isolates of multinucleated *Rhizoctonia* solani AG-2.

Isolates	Nuclear condition	Identification	Isolated in	Geographical origin
J1	Multinucleate L-9-1S	Thanatephorus cucumeris	Zoysia grass	Gifu, Japan
J2	Multinucleate G4	Thanatephorus cucumeris	Zoysia grass	Gifu, Japan
J3	Multinucleate AJ-1-10-1	Thanatephorus cucumeris	Zoysia grass	
J4	Multinucleate RGR38	Thanatephorus cucumeris	Zoysia grass	Shizuoka, Japan
J5	Multinucleate RGR 39	Thanatephorus cucumeris	Zoysia grass	Shizuoka, Japan
J6	Multinucleate SLP3-3	Thanatephorus cucumeris	St. Augustine grass	Okinawa, Japan
J7	Multinucleate RGR-59	Thanatephorus cucumeris	Bent grass	Shizuoka, Japan
J8	Multinucleate C-116	Thanatephorus cucumeris	Mat rush	Kumamoto, Japan
J9	Multinucleate C-306	Thanatephorus cucumeris	Mat rush	
J10	Multinucleate 9207	Thanatephorus cucumeris	Beefsteak plant	Hyogo, Japan
J11	Multinucleate No. 2	Thanatephorus cucumeris	Maize	Miyazaki, Japan
J12	Multinucleate Gu-1	Thanatephorus cucumeris	Guzmania	Gifu, Japan

J13	Multinucleate K-1	Thanatephorus cucumeris	Sugar beet	Hokkaideo. Japan
J14	Multinucleate S-2	Thanatephorus cucumeris	Sugar beet	Hokkaido. Japan
J15	Multinucleate R56	Thanatephorus cucumeris	Sugar beet	Hokkaido. Japan
J16	Multinucleate R64	Thanatephorus cucumeris	Sugar beet	Hokkaido. Japan
J17	Multinucleate 92155-1-2	Thanatephorus cucumeris	Burdock	Hokkaido. Japan
J18	Multinucleate 92155-1-3	Thanatephorus cucumeris	Burdock	Hokkaido. Japan

Pathogenicity assays

Pathogenicity tests were performed with the strains isolated from the field in Albacete on the following plant species: radish (*Raphanus sativus*), cauliflower (*Brassica oleracea* var. *botrytis*), onion (*Allium cepa*), turf grasses: zoysia grass (*Zoysia tenuifolia*), St. Augustine grass (*Stenotaphrum secundatum*) and Bermuda grass (*Cynodon dactylon*), tomato (*Lycopersicum esculentum*), carrot (*Daucus carota*) and cucumber (*Cucumis sativum*). Seeds were surface-disinfected in 1% sodium hypochlorite for 10 min, rinsed with sterile distilled water and aseptically blotted between two layers of wet sterile paper into petri dishes. They were allowed to germinate for 4 days at 260 C in the dark. Isolates were grown on water agar for 2 days at 260 C. Five germinated seeds were then placed around the fungal colony. The plates were incubated at room temperature in the lab bench, and pathogenicity was assayed for the first time 7-10 days after the tests were set, and subsequent observations were made afterwards every two days. The *Rhizoctonia* solani AG 4 strain Me8-2 (Boysen *et al.*, 1996) was used as positive control, since this strain is known to be very pathogenic to all plant species tested. A negative control of germinated seeds without fungus was included in each of the tests.

Protection assays

The strains that did not show pathogenicity towards one or more plant species were tested for protection from the pathogenic strain Me 8.2. This was done by inoculating both the pathogenic strain and the supposedly protective strain. This was done in two ways, simultaneously (both strains inoculated at the same time) and successively (first the supposedly protective strain was allowed to colonise the plants surface for three days and then the pathogenic strain was inoculated). A negative control (no fungus), a positive control (the pathogenic strain Me 8.2), and the supposedly protective strain alone were included for each of the plant species in the tests. Protection was assayed for the first time 7 days after the tests were set, and subsequent observations were made every two days.

Isolation of genomic DNA, asymmetric PCR and DNA sequencing procedures

For DNA extraction, *Rhizoctonia* plugs were inoculated in 50 ml complete medium [CM: 5 g malt extract (Merck, Darmstadt, Germany), 5 g yeast extract (Difco, Detroit, USA) and 5 g glucose anhydrous (Merck, Darmstadt, Germany) per litre] in a flask, the mycelium was grown for 4 days at 23° C and 150 rpm agitation, blotted dry and freeze dried. After grinding, mycelium was suspended in 500 ml DNA-extraction buffer [20 mM Tris/HCl pH 8.0, 50 mM EDTA and 0.4 % (w/v) SDS], incubated at room temperature for 5 min and extracted twice with an equal volume of phenol: chloroform: isoamylalcohol (25:24:1). Subsequently, the aqueous phase was incubated for 30 min at 37° C in the presence of 50 mg/ml RNase A and then extracted with chloroform: isoamylalcohol (24:1). DNA was precipitated by adding 0.6 volumes isopropanol, washed with 70% ethanol, dried and resuspended in 100 ml ultrapure sterile water.

Eucariotical ribosomal genes are arranged in a tandem repeat with the 5.8S gene flanked by internal transcribed spacers (ITS). Fig 1 is a schematic representation of the Internal Transcribed Spacer (ITS) region of the ribosomal DNA. The shadowed boxes represent the ribosomal genes. The arrows denote the position of the PCR and sequencing primers.

The regions of rDNA repeat from the 3' end of the 18S gene to the 5' end of the 28S gene were amplified using the primers ITS 5, 4 , 4B and 1F (White et al, Gardes et al). Asymmetric PCR to synthesise ssDNA of the ITS region was done with a 50:1 molar ratio between two primers (Gyllenstein). Fragments of approximately 700 bp were obtained when using *Rhizoctonia*

genomic DNA as a template for PCR amplification of the ITS region. Different combinations of primers were tested for the asymmetric amplification. Asymmetric PCR products using ITS4/ITS5 in the ratio 50:1 were found to give better results when sequencing with ITS1 and ITS3, whereas the combination of ITS1F/ITS4B (also in the 50:1 ratio) was found to give better results when sequencing with ITS2 and ITS4. Manual sequencing was performed using the Sanger's dideoxy method (Sanger *et al.*, 1977).

Phylogenetic analysis

The sequence data were checked between complementary strands using GCG Fragment Assembly System (Program Manual for the Wisconsin Package, version 8, Wisconsin). The resulting sequences were aligned using the software package CLUSTALW (Thompson *et al.*, 1994) (IntelliGenetics, Inc., Mountain View, CA). All the data were processed for phylogenetic analysis of the nucleotide and gap differences. The distances in the ITS-5.8S rDNA region were determined by Kimura two parameter model (Kimura, 1980), omitting all sites with gaps. Phylogenetic trees were constructed using the maximum-parsimony analysis with the heuristic search of the Phylogeny Using Parsominy Analysis (PAUP) program 3.1.1. (Swofford, 1991) which gaps treated as missing data. The branch and bound algorithm provides an efficient approach to find all the minimal lengths. The trees were rooted with the rDNA sequence of Trichoderma harzianum. A bootstrap analysis using 1000 resample of the sequence data was carried out, and using the heuristic search option of PAUP. Trees were also obtained by using the Neighbour-joining method (Saitou and M., 1987) as it is in the program "Neighbour" from the PHYLIP package version 3.5 (Felsenstein, 1993). The topological accuracy was again estimated by the bootstrap method with 1000 replicates. The percentage of bootstrap replicates which yields each grouping as a measure of statistical confidence. A grouping found on 95% of bootstrap replicates was considered statistically significant.

RESULTS

Pathogenicity assays

Crucifers (radish and cauliflower) are susceptible to all *Rhizoctonia* isolates tested, regardless of whether they were multinucleated or binucleated. The same was true for onion. Grass, tomato and carrot were susceptible to all multinucleated *Rhizoctonia*e (Me 8.2, aA1, aA2), but were not susceptible to any of the binucleate strains, therefore protection tests were done with these plant species. The strains randomly chosen to test protection from the pathogenic strain Me 8.2 were the binucleate strains aB and F2. Cucumber was susceptible to all strains tested, except for S7, therefore this strain was subsequently tested for protection from the pathogenic strain Me 8.2. The strain Me 8.2 showed significatively stronger pathogenicity than any other strains tested, and this was seen in shorter times needed to attack the plant, and complete killing of the seedling, whereas with other pathogenic strains the seedlings were damaged, but not as much affected as with isolate Me 8.2. In the pathogenicity tests was also observed that in some cases the seeds had not been completely surface sterilised, and in the negative control (no fungus) was observed growth of some fungi other than *Rhizoctonia*, yeast or bacteria, but when *Rhizoctonia* had colonised the seedlings, there was no growth of these contaminants.

Protection essays

For grass, tomato and carrot, it was observed that the negative control (no fungus) presented contamination from other fungus, yeast and/or bacteria, whereas the control with protective strains B and F2 were clean, that is, *Rhizoctonia* did not allow any of the contaminants carried in the seed to grow. As expected, plants inoculated with protective strains B and F2 looked completely healthy, whereas plants inoculated with Me 8.2 were heavily injured. In the simultaneous test (that is, protective and pathogenic strains inoculated at the same time) there was no protection of the binucleate strains from the pathogenic Me 8.2. But in the successive test, in which the binucleate strains B and F2 were allowed to colonise the plant surface prior to adding the pathogenic strain, a very clear protective effect was observed, and the plants looked as healthy as the negative control and the control with the binucleate strains alone (Fig. 3). Furthermore, it was also observed that the seedlings covered by the protective strains looked slightly bigger than the negative control, thus pointing to the possibility that there is an increased growth response induced by the non-pathogenic *Rhizoctonia* isolates (Fig. 3). This has already been described by (Sneh *et al.*, 1986). In cucumber, although some protection was seen, we could not observe a clear

protective effect of strain S7, therefore we could not conclude that strain S7 acts as a protective strain in cucumber.

We foresee the same protective behaviour foe the entire group of the new species including F2. B protects much better than A1 and A2 which are pathogenic.

Sequencing PCR products from the PCR amplified ITS regions

PCR experiments carried out on the ITS-5.8S rDNA region of 20 isolates of binucleate *Rhizoctonia* spp (including isolates of Ceratobasidium from CBS collection and RBN testers) A single band was detected in each amplification reaction when analysed by agarose gel electrophoresis. The amplified DNA was sequenced directly in both strands.

Primers obtained from the sequenced isolates of binucleated *Rhizoctonia* for their identification

Albacete

Ab -1

CCCTATTAAGGGGCA

Binucleated

Bin

CCGTAAAAAAGTCTT

Cerealis

Cer11

CGCGAGAGAGAGGCTGGC

Albacete

Ab+

CCACACAAACCCATTGTATT

Cornigerum

Corg

AATTGTTCCTTTGGGCAC

Albacete

Ab-2

GGAACCTTTATTGGAC

Cerealis

Cerl2

AATGTAATCGATGTAAACGC

Albacete

Ab-3

CTCAATTTATTTTAAAACGA

The molecular definition of the members of the new species group may be based on several specific sequences of the ITS segment. The sequence CCACACAAACCCATTGTATT of the primer AB+ is common and specific to the group of F2, with the exception of S6 that changes one nucleotide (ATCG instead of ATTG), see pages 3-4 of the manual alignment.

This primer may define the group whereas we cannot define the isolate B. This isolate has a unique property of having cells with two and with four or more nuclei. This characteristic, never described before may be important for the understanding of the relationship between binucleated and multinucleated *Rhizoctonia*.

Phylogenetic analysis

The CLUSTALW sequence alignments are shown in Exhibit A.

Reference is also made to the phylogenetic tree of Figure 2. The 5.8S rDNA gene sequence (155 bp) was conserved among the *Rhizoctonia* isolates examined. Among the isolates analysed, the sequence of ITS 1 was more divergent than that of the ITS 2 region. There was 100-81%

ˏequence homology among the sequences of total ITS regions of all the isolates used in this study. Sequence homology among binucleate isolates varied from 99-85%, and 85-75% between binucleate and multinucleated isolates (AG 2). In phylogenetic analysis a consensus sequence of the ITS regions of 18 AG 2 isolates was used. Phylogenetic analysis separates binucleate and multinucleated in two different clusters. In the multinucleated cluster it is possible to differentiate AG 2 and AG 4, and also subgroups within each AG that correlate with pathogenicity and geographical origin. Binucleate isolates clustered together in an independent branch. It is interesting that Spanish isolates from Albacete cluster differentiated from binucleate isolates from the CBS collection. Multinucleated isolates from alfalfa (A1 and A2) cluster in AG 4 group, with similar features of pathogenicity. In the tree, horizontal lengths in the branches indicate the relative genetic distances between the isolates.

F1, F6, F5, F7, S5, S6, F2, F3 and S1 thus comprise the new species so far. S3 clusters with reference strains and should be similar to AGN, AGP, AGBO, and AGK. S2 and S$ are similar to AGL, AGI, AGC and AGD.

REFERENCES

Boysen, M., Borja, M., del Moral, C., Salazar, O. and Rubio, V. (1996) Identification at strain level of *Rhizoctonia* solani AG4 isolates by direct sequence of asymmetric PCR products of the ITS regions Curr. Genet., 29, 174-181.

Cubeta, M. A., Echandi, E., Abernethy, T. and Vilgalys, R. (1991) Characterization of anastomosis groups of binucleate *Rhizoctonia* species using restriction analysis of an amplified ribosomal RNA gene Phytopathology, 81, 1395-1400.

Duncan, S., Barton, J. E. and O'Brien, P. A. (1993) Analysis of variation in isolates of *Rhizoctonia* solani by random amplified polymorphic DNA assay Mycol. Res., 97, 1075-1082.

Felsenstein, J. (1993) PHYLIP manual version 3.5. Washington, University of Washington.

'abaji-Hare, S. H., Meller, Y., Gill, S. and Charest, P. M. (1990) Investigation of genetic relatedness among anastomosis groups of *Rhizoctonia* solani using cloned DNA probes Can. J. Plant Pathol., 12, 393-404.

Julian, M., Dullemans, A., van Silfhout, C. H. and Keijer, J. (1997) Nuclear behaviour in homokaryotic and heterokariotic fruiting of Thanatephorus cucumeris (*Rhizoctonia* solani) anastomosis group 1, subgroup IC Mycologia, 89, 361-374.

Kataria, H. R. and Gisi, U. (1989) Recovery from soil and sensitivity to fungicide of *Rhizoctonia* solani Mycol. Res., 92, 458-462.

Keijer, J., Houterman, P. M., Dullemans, A. M. and Korsman, M. G. (1996) Heterogeneity in electrophoretic karyotype within and between anastomosis groups of *Rhizoctonia* solani Mycol. Res., 100, 789-797.

Kimura, M. (1980) A simple method for estimating evolutionary rates through comparative studies of nucleotide sequences. J. Mol. Evol, 16, 111-120.

Kuninaga, S. and Yokosawa, R. (1980) A comparison of DNA base compositions among anastomosis groups in *Rhizoctonia* solani Kühn Ann. Phytopath. Soc. Japan, 46, 150-158.

Kuninaga, S. and Yokosawa, R. (1982) DNA base sequence homology in *Rhizoctonia* solani Kuhn II. Genetic relatedness within anastomosis group 2 Ann. Phytopath. Soc. Japan, 48, 668-673.

Kuninaga, S. and Yokosawa, R. (1982) DNA base sequence homology in *Rhizoctonia* solani Kuhn. I. Genetic relatedness within anastomosis group 1 Annu. Phytopath. Soc. Japan, 48, 659-667.

Kuninaga, S. and Yokosawa, R. (1984) DNA base sequence homology in *Rhizoctonia* solani Kuhn. IV. Genetic relatedness within AG-4 Annals. Phytopathol. Soc. Japan, 50, 322-330.

Kuninaga, S. and Yokosawa, R. (1985) DNA base sequence homology in *Rhizoctonia* solani Kuhn. V. Genetic relatedness within AG-6 Annu. Phytopathol. Soc. Japan, 50, 346-352.

Kuninaga, S. and Yokosawa, R. (1985) DNA base sequence homology in *Rhizoctonia* solani Kuhn. VI. Genetic relatedness among seven anastomosis groups Annals Phytopathological Society of Japan, 51, 127-132.

Kuninaga, S. and Yokosawa, R. (1985) DNA base sequence homology in *Rhizoctonia* solani Kuhn. VII. Genetic relatedness between AG-BI and other anastomosis groups Annu. Phytopathol. Soc. Japan, 51, 133-138.

Saitou, N. and M., Nei. (1987) The neighbour-joining method: a new method for reconstructing phylogenetic trees Molecular Biology and Evolution, 4, 406-425.

Salazar, O., Schneider, H., Keijer, J. and Rubio, V. (1997) Phylogenetic relations in *Rhizoctonia* solani AG2 based on ribosomal internal transcribed spacers (ITS) sequences Ready for submission

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) DNA sequencing with chain-termination inhibitors Proc. Natl. Acad. Sci. USA, 74, 5463-5467.

Sneh, B., Zeidan, M., Ichielevich-Auster, M., Barash, I. and Koltin, Y. (1986) Increased growth responses induced by a nonpathogenic *Rhizoctonia* solani Can. J. Bot., 64, 2372-2378.

Swofford, D. L. (1991) PAUP: Phylogenetic Analysis Using Parsimony, version 3.1.1. Champaign, IL., Illinois Natural History Survey.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice Nucl. Acids Res., 22, 4673-4680.

Vilgalys, R. and Gonzalez, D. (1990) Ribosomal DNA restriction fragment length polymorphisms in *Rhizoctonia* solani Phytopathology, 80, 151-158.

17

EXHIBIT A

C JSTALW ALIGNMENT

C116	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTTGGTTG
RGR-59	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTAGGTTG
9207	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTTGGTTG
NO	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTAGGTTG
C306	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTAGGTTG
GU1	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTTGGTTG
AJ1101	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTAGGTTG
G4	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTAGGTTG
RGR39	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTAGGTTG
L91S	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTAGGTTG
RGR38	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTAGGTTG
SLP33	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTAGGTTG
9215512	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTAGGTTG
K1	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTAGGTTG
9215513	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTAGGTTG
Sdos	--------------------	--------------------	--------------------	---------TAAACAAAGG-	--------GTATTTAGGTTG
R64	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTATTTAGGTTG
R56	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTGTTTAGGTTG
R65	--------------------	--------------------	--------------------	--------TTAAACAAAGG-	--------GTGTTTAGGTTG
S5	-------------------T	GGATAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATG--	AAT--ATAGAG--TTGGTTG
S6	------------------GG	GATAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATG--	AAT--ATAGAG--TTGGTTG
F1	----------TTAGAGGGGA	ATAAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATG--	AAT--ATAGAG--TTGGTTG
F2	--------------AGAGGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATG--	AAT--ATAGAG--TTGGTTG
F3	------------TTAGAGGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATG--	AAT--ATAGAG--TTGGTTG
F5	-----------------GGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATG--	AAT--ATAGAG--TTGGTTG
F7	---TTCCCTTTTTTAGAGGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATG--	AAT--ATAGAG--TTGGTTG
F6	---------TTTAGAGGGAA	GTAAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATG--	AAT--ATAGAG--TTGGTTG
S1	--------------------	------AGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATG--	AAT--ATAGAG--TTGGTTG
A2	--------------------	----AAAGTCG-AACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATT--	AAT--GTAGAGT-TTGGTTG
A1	--------------------	----AAAGTCG-AACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATT--	AAT--GTAGAGT-TTGGTTG
B	--------------------	-----AAGTCG-AACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATT--	AAT--GTAGAGT-TTGGTTG
AGK	---CTTGGTCATTTAGAGGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAACCAC	-----TTGGAGTC--GGTTG
AGN	---CTTGGTCATTTAGAGGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATGA-	AA---CTAGAGTT--GGTTG
AGP	---CTTGGTCATTTAGAGGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATG--	AACACTTGGAGTC--GGTTG
AGBO	GCTCTTGGTCATTTAGAGGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATG--	AAC--TTAGAGTC--GGTTG
AGBB	GCTCTTGCTCATTTAGAGGA	AGTAAAAGTCGTAAGAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATG--	AAC--TTAGAGTC--GGTTG
S3	--------------------	----AAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATG--	AAC--TTAGAG--TCGGTTG
AGJ	---CTTGGTCATTTAGAGGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATGA-	AA---CTAGAGTT--GGTTG
AGG	--TCTTGGTCATTTAGAGGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATG--	AAT--GTAGAGTT--GGTTG
C5	--------------------	------------AACAAGGT	TTCCGTAGGTGCACCTGCGG	AAGGATCATTAATGAATGT-	AAT--GTGGAGT-T-GGTTG
C8	--------------------	----------CTAACAAGGT	TTCCGTAGGTGAACCTGCGG	ATGGATCATTAATGAATTT-	AAT--GTGGAGT-T-GGTTG
C2	--------------------	-------GATGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATTTT	AAT--GAGGAGT-TTGGTTG
C4	--------------------	---------TGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATTTT	AAT--GAGGAGT-T-GGTTG
C6	--------------------	--------ATGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATTT-	AAT--GTAGAGT-T-GGTTG
C1	--------------------	-------GATGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATTTT	AAT--GAGGAGT-TTGGTTG
C11	-------------------A	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTAATGAATG--	AAT--GTAGAGT-C-GGTTG
C12	-------------------A	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTAATGAATG--	AAT--GTAGAGT-C-GGTTG
C13	-------------------A	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTAATGAATG--	AAT--GTAGAGT-C-GGTTG
S2	--------------------	-------GTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATCGAATG--	AAC--AATGGGGATCGGTTG
S4	--------------------	-------GTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATCGAATG--	AAC--AATGGGGATCGGTTG
AGL	--TCTTGGTCATTTAGAGGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATG--	AAC--ATGGAGT-TTGGTTG
AGI	--TCTTGGTCATTTAGAGGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATGA-	AA---CTGGAGT-TTGGTTG
AGC	---CTTGGTCATTTAGAGGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATGA-	AA---CTAGAGTT--GGTTG
AGA	---CTTGGTCATTTAGAGGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATGA-	AA---CTAGAGTT--GGTTG
CAG2	---CTTGGTCATTTAGAGGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATGA-	AA---CTAGAGTT--GGTTG
AGD	--TCTTGGTCATTTAGAGGA	AGTAAAAGTCGTAACAAGGT	TTCCGTAGGTGAACCTGCGG	AAGGATCATTATTGAATGA-	AA---CTAGAGTT--GGTTG
AGE	---CTTGGTCATTTAGAGGA	AGTATTAGTCGTAACAAGGT	TTCCGTAGGTAAACCTGCGG	AAGGATCATTATTGAATGA-	AA---CTAGAGTT--GGTTG
AGR	--------------------	--------------------	---------------CGCGG	A--GGTC-TTAATGAATGA-	AA----GGGAGTT------G
AGS	--------------------	--------------------	---------------CGCGG	A--GGTC-TTAATGAATGA-	AA----GGGAGTT------G
THARZIANU	--------------------	--TAGAAGTCGTAACAAGGT	CTCCGTTGGTGAACCAGCGG	AGGGATCATTACCGAGTT--	------TACAA--------C

```
C116      TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
-59       TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
207       TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATGA  ---TGCACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
NO        TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
C306      TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
GU1       TAGCTGGCTC---CATTA--  ------ATTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
AJ1101    TAGCTGGCTC---CGTTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
G4        TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
RGR39     TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
L91S      TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
RGR38     TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
SLP33     TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
9215512   TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
K1        TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
9215513   TAGCTGTCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
Sdos      TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGGACACCTTTTGCTCT  TTTTTT-AATCC--ACACAC  --ACCTGTGAACCTGTG-AG
R64       TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTTTAATCC--ACACAC  --CCCTGTGAACCTGTG-AG
R56       TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTTTAATCC--ACACAC  --ACCTGTGAACCTGTG-AG
R65       TAGCTGGCTC---CATTA--  ------GTTTGGA-GCATG-  ---TGCACACCTTTTGCTCT  TTTTTTTAATCC--ACACAC  --ACCTGTGAACCTGTG-AG
S5        TCGCTGGCT----CCT----  ------C-CGGGA-GCAT--  --GTGCACGC-TTT--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACTTGTG-AG
S6        TCGCTGGCT----CCT----  ------C-CGGGA-GCAT--  --GTGCACGC-TTT--CTCT  --TTC---ATCC--ACACAC  --ACGTGTGCACTTGTG-AG
F1        TCGCTGGCT----CCT----  ------C-CGGGA-GCAT--  --GTGCACGC-TTT--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACTTGTG-AG
F2        TCGCTGGCT----CCT----  ------C-CGGGA-GCAT--  --GTGCACGC-TTT--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACTTGTG-AG
F3        TCGCTGGCT----CCT----  ------C-CGGGA-GCAT--  --GTGCACGC-TTT--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACTTGTG-AG
F5        TCGCTGGCT----CCT----  ------C-CGGGA-GCAT--  --GTGCACGC-TTT--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACTTGTG-AG
F7        TCGCTGGCT----CCT----  ------C-CGGGA-GCAT--  --GTGCACGC-TTT--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACTTGTG-AG
F6        TCGCTGGCT----CCT----  ------C-CGGGA-GCAT--  --GTGCACAC-TTT--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACTTGTG-AG
S1        TCGCTGGCT----CCT----  ------CGCGGGA-GCAT--  --GTGCACGC-TTT--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACTTGTG-AG
A2        TAGCTGGCT----CCTAATT  ---AAACTTGGGG-GCAT--  --GTGCACACCTTT--CTCT  --TTC---ATCCC-ATACAC  --ACCTGTGCACCTGTG-AG
A1        TAGCTGGCT----CCTAATT  ---AAACTTGGGG-GCAT--  --GTGCACACCTTT--CTCT  --TTC---ATCCC-ATACAC  --ACCTGTGCACCTGTG-AG
B         TAGCTGGCT----CCTAATT  ---AAACTTGGGG-GCAT--  --GTGCACACCTTT--CTCT  --TTC---ATCCC-ATACAC  --ACCTGTGCACCTGTG-AG
AGK       TCGCTGGCTG---TCT---T  -----ATTTGGCA-GCAT--  --GTGCACGCCT-T--CTCT  T-TTC---ATCCACACACAC  --CCCTGTGCACTTGTG-AG
AGN       TCGCTGGCTG---TCTT---  --------TGGCA-GCAT--  --GTGCACGCCT-T--CTCT  T-TTC---ATCCACACAC--  --CCCTGTGCACTTGTG-AG
AGP       TCGCTGGCTG---TCTT---  --------TGGCA-GCAT--  --GTGCACGCCT-T--CTCT  A-TTC---ATCC--ACACAC  --CCCTGTGCACCTGTG-AG
AGBO      TCGCTGGCTG---TCTT---  --------TGGCA-GCAT--  --GTGCACGCCT-T--CTCT  A-TTC---ATCC--ACACAC  --CCCTGTGCACCTGTG-AG
AGBB      TCGCTGGCTG---TCTT---  --------TGGCA-GCAT--  --GTGCACGCCT-T--CTCT  A-TTC---ATCC--ACACAC  --CCCTGTGCACTTGTG-AG
S3        TCGCTGGCT-----GTC---  ------TTTGGCA-GCAT--  --GTGCACGC-CTT--CTCT  A-TTC---ATCC--ACACAC  --ACCTGTGCACTTGTG-AG
AGJ       TCGCTGGCCC----CT----  T----CACTGGG--GCAT--  --GTGCACGCCT-T--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACTTGTG-AG
AGG       TCGCTGGCCC----TT----  T-----CACKGGG-GTAT--  --GTGCACGCCT-T--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGAACTTGTG-AG
C5        TAGCTGGTC-----CTAATT  ----TCCTTTGG--ACAT--  --GTGCACACCT-T--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGAACTAGTG-AG
C8        TAGCTGGTC-----CTAATT  ----TCCTTTGGG-TCAT--  --GTGCACACCT-T--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACTTGTG-AG
C2        TAGCTGGCC-----CTGATT  G-TTTCCTTTGG--CAATTT  -TGTGCACGCCTCT--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACTTGTG-AG
C4        TTGCTGGCC-----TTAATT  G-TTCCA---GG---CACT-  -TGTGCACACCT-T--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGAACCTGTG-AG
C6        TAGCTGGC-------TT---  ---TTCTGTGGGG-GCG---  -GGTGCACACCT-T--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACTTGTG-AG
C1        TAGCTGGCC-----CTGATT  G-TTTCCTTTGG---CAATT  TTGTGCACGCCTCT--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACTTGTG-AG
C11       TAGCTGGGT-----CT---T  TTAATCGA--GG--CCAT--  --GTGCACGCCT-T--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACCTGTTTAG
C12       TAGCTGGGT-----CT---T  TTAATCGA--GG--CCAT--  --GTGCACACCT-T--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACCTGTTTAG
C13       TAGCTGGGT-----CT---T  TTAATCGA--GG--CCAT--  --GTGCACGCCT-T--CTCT  --TTC---ATCC--ACACAC  --ACCTGTGCACCTGTTTAG
S2        TCGCTGGCTT----CTT---  ------CATTGAG-GCAT--  --GTGCACGCCTTT--CTCT  ATTTC---ATCC--ACACAC  --ACCTGTGAACTTGTG-AG
S4        TCGCTGGCTT----CTT---  ------CATTGAG-GCAT--  --GTGCACGCCTTT--CTCT  ATTTC---ATCC--ACACAC  --ACCTGTGAACTTGTG-AG
AGL       TTGCTGGCCC----CT----  T----TATTGGG--GCAT--  --GTGCACACCT-C--CTCT  A-TTC---ATCC--ACACAC  --ACCTGTGAACTTGTG-AG
AGI       TTGCTGGCCC----CT----  T----TATTGGG--GCAT--  --GTGCACACCT-C--CTCT  A-TTC---ATCC--ACACAC  --ACCTGTGAACTTGTG-AG
AGC       TCGCTGGCCC----CT----  T----CACTGGG--GCAT--  --GTGCACGCCT-C--CTCT  A-TTC---ATCC--ACACAC  --ACCTGTGAACTCCCG-AG
AGA       TCGCTGGCCC----CT----  T----CACTGGG--GCAT--  --GTGCACGCCT-C--CTCT  A-TTC---ATCC--ACACAC  --ACCTGTGAACTCCCG-AG
CAG2      TCGCTGGCCC----CT----  T----CACTGGG--GCAT--  --GTGCACCCCT-C--CTCT  A-TTC---ATCC--ACACAC  --ACCTGTGAACTCCCG-AG
AGD       TCGCTGGCCC----CT----  T----CACTGGG--GCAT--  --GTGCACGCCT-C--CTCT  A-TTC---ATCC--ACACAC  --ACCTGTGAACTTGTG-AG
AGE       TCGCTGGCCC----CT----  T-----CACTGGGGCCAT--  --GTGCACGCCT-C--CTCT  A-TTC---ATCC--ACACAC  --ACCTGTGAACTTGTG-AG
AGR       ATGCTGGCCA----------  ---GGTCTCTGGC---AT--  --GTGCTCGCTTCCGACAAT  C-TT----ATCC--ATAC--  --ACCTGTGCAC--------
AGS       ATGCTGGCCA----------  ---GGTCTCTGGC---AT--  --GTGCTCGCTTCCGACAAT  C-TT----ATCC--ATAC--  --ACCTGTGCAC--------
THARZIANU TCCCAAAC-----CC-----  ------AATGTGAACCAT--  ---ACCAAACTGTTGCCTCG  ---GCGGGGTCACGCCCCGG  GTGCGTCGCAGCCCCGG-AA
```

```
      5  GCAGAGG---CATGGATGGG AG----AACTTTCATTTACC TAAAAA-TAAATGTATTGGG ACCCCCATCCCCCC--TCTG TCTACTCAACT-CTAA----
  RG. 59 GCAAAGA---CATGGATGGG AG----AACTTTCATTTACC TAAAAA-TAAATGTATTGGG ACCCCCATCCCCCC--TCTG TCTACTCAACT-CTAA----
    9207 GCAGAGG---CATGGATGGG AG----AACTTTCATTTACC TAAAAA-TAAATGTATTGGG ACCCCCATCCCCCC--TCTG TCTACTCAACT-CTAA----
      NO GCAGAGA---CATGGATGGG AG----AACTTTCATTTACC TAAAAA-TAAATGTATTGGG ACCCCCATCCCCCC--TCTG TCTACTCAACT-CTAA----
    C306 GCAGAGA---CATGGATGGG AG----AACTTTCATTTACT TTAAAA-TGAATG-ATTGGG ACCCCCATCCCCCC--TCTG TCTACTCAACT-CTAA----
     GU1 GCAGAGA---CATGGATGGG AG----AACTTTCATTTACT TAAAAA-TAAATGTATTGGG AACCCCAACCCCCC--TCTG TCTACTCAACT-CTAA----
  AJ1101 GCAAAGAAA-CATGGATGGG CA----AACTTTCATTTACT TTTAAA-TAAGTG-ATTGGG ACCCCTACCCCCCCC-TCTG TCTACTCAACT-CTAA----
      G4 GCAAAGAAA-CATGGATGGG CA----AACTTTCATTTACT TTTAAA-TAAGTG-ATTGGG ACCCCTACCCCCCC--TCTG TCTACTCAACT-CTAA----
   RGR39 GCAAAGAAA-CATGGATGGG CA----AACTTTCATTTACT TTTAAA-TAAGTG-ATTGGG ACCCCTACCCCCCC--TCTG TCTACTCAACT-CTAA----
    L91S GCAAAGAAA-CATGGATGGG CA----AACTTTCATTTACT TTTAAA-TAAGTG-ATTGGG ACCCCTACCCCCCCC-TCTG TCTACTCAACT-CTAA----
   RGR38 GCAGAGAAA-CATGGATGGG CA----AACTTTCATTTACT TTTAAA-TAAGTG-ATTGGG ACCCCTACCCTCCCC-TCTG TCTACTCAACT-CTAA----
   SLP33 GCAGAGAAA-CATGGATGGG CA----AACTTTCATTTACT TTTAAA-TAAGTG-ATTGGG ACCCCTACCCTCCCC-TCTG TCTACTCAACT-CTAA----
 9215512 GCAGAGA---CATGGATGGG AA----AACTTTTGTTTACT TTTAGG-TGAATG-ATTGGG ACCCCTACCCCCC---TCTG TCTACTCAACT-CTAA----
      K1 GCAGAGA---CATGGATGGG AA----AACTTTTGTTTACT TTTAGG-TGAATG-ATTGGG ACCCCTACCCCCC---TCTG TCTACTCAACT-CTAA----
 9215513 GCAGAGA---CATGGATGGG AA----ACCTTTTGTTTACT TTTAGG-TGAATG-ATTGGG ACCCCTACCCCCC---TCTG TCTACTCAACT-CTAA----
    Sdos GCAGAGA---CATGGATGGG AA----AACTTTTGTTTACT TTTAGG-TGAATG-ATTGGG ACCCCTACCCCCC---TCTG TCTACTCAACT-CTAA----
     R64 GCAGAGA---CATGGATGGG AA----AACTTTCATTTACT TTTAGG-TGAATG-ATTGGG ACCCCTACCCCCC---TCTG TCTACTCAACT-CTAA----
     R56 GCAGAGA---CATGGATGGG AA----AACTTTTGTTTACT TTTAGGGTAAATG-ATTGGG ACCCCTACCCCCCC--TCTG TCTACTCAACT-CTAA----
     R65 GCAGAGA---CATGGATGGG AA----AACTTTTGTTTACT TTTAGGGTAAATG-ATTGGG ACCCCTACCCCCCC--TCTG TCTACTCAACT-CTAA----
      S5 ACGGA--------GGACCGT AA-----AAAA--------- -------------------- ------GTCTTCC------G TCTATTAAAC----CA----
      S6 ACGGA--------GGACCGT AA-----AAAA--------- -------------------- ------GTCTTCC------G TCTATTAAAC----CA----
      F1 ACGGA--------GGACCGT AA-----AAAA--------- -------------------- ------GTCTTCC------G TCTATTAAAC----CA----
      F2 ACGGA--------GGACCGT AA-----AAAA--------- -------------------- ------GTCTTCC------G TCTATTAAAC----CA----
      F3 ACGGA--------GGACCGT AA-----AAAA--------- -------------------- ------GTCTTCC------G TCTATTAAAC----CA----
      F5 ACGGA--------GGACCGT AA-----AAAT--------- -------------------- ------GTCTTCC------G TCTATTAAAC----CA----
      F7 ACGGA--------GGACCGT AA-----AAAA--------- -------------------- ------GTCTTCC------G TCTATTAAAC----CA----
      F6 ACGGA--------GGACCGT AA-----AAAA--------- -------------------- ------GTCTTCC------G TCTATTAAAC----CA----
      S1 ACGGA--------GGACCGT AA-----AAAA--------- -------------------- ------GTCTTCC------G TCTATTAAAC----CA----
      A2 ACAGATGTTTTGTGGGGGGG AA--GGAACTTTATTGGACC TTCTA--------------- ---CTCCCCCTTGACTTCTG TCTACTTAATT---CA----
      A1 ACAGATGTTTTGTGGGGGGG AA--GGAACTTTATTGGACC TTC----TA----------- ---CTCCCCCTTGACTTCTG TCTACTTAATT---CA----
       B ACAGATGTTTTGTGGGGGGG AA--GGAACTTTATTGGACC TTC----TA----------- ---CTCCCCCTTGACTTCTG TCTACTTAATT---CA----
     AGK ACT--------GGAGACCGG TAA---AAAAA--------- -------------------- ------GTCTTCA------G TCTGCTAAATC---CA----
     AGN ACTA---------GGAGGCC GT----AAAAA--------- -------------------- -------GCTTTCA-----G TCTGCTAAATT---CA----
     AGP ACT----------GGAGGCC GT----AAAAA--------- -------------------- ------GCCTTCA------G TCTGCTAAATT---CA----
    AGBO ACT----------GGAGCCC GT----AAAAA--------- -------------------- ------GCCTTCA------G TCTGCTAAATT---CA----
    AGBB ACT----------GGAGCCC CT----AAAAA--------- -------------------- ------GCCTTCA------G TCTGCTAAATT---CA----
      S3 ACTGG--------AGGCCGT AA----AAA----------- -------------------- ------GCCTTCA------G TCTGCTAAATT---CA----
     AGJ AC-----------GGAGGGC TTT----AATTA-------- -------------------- ------GTCTTCC------G TCTACTTAAT----C-----
     AGG AC-----------GGAGGGC TTT----AATTA-------- -------------------- ------GTCTTCC------G TCTATTCAAC----C-----
      C5 GCAGA--------GGACCGT AA----AAAAA-----GTCT T------------------- -----------------CTG TCTACTTAA---TTTA----
      C8 GCAGA--------GGACCGT AA----AAAA------GTCT T------------------- -----------------CTG TCTACTTAA---TTTA----
      C2 ACAGA--------GGACCGT AA----AAAA------GTCT T------------------- -----------------CTG TCTACTTAA---TCCA----
      C4 ACAGA--------GGACCGT AA----AAAA------GTCT T------------------- -----------------CTG TCTACTTAA---TTTA----
      C6 ACAGA--------GGAC-GT AA----AAAA------GTCT T------------------- -----------------CTG TCTACTTAA---TTCA----
      C1 ACAGA--------GGACCGT AA----AAAA------GTCT T------------------- -----------------CTG TCTACTTAA---TCCA----
     C11 ACGGT--------CGAAGGA AA----AAGTCTTTCTCGCG AGAGAGAGG----------- -----------------CTG GCTCCTTTTCCGTCCAA---
     C12 ACGGT--------TGAAGGA AA----AAGTCTATCTCGCG AGAGAGAGG----------- -----------------CTG GCTCCTTTTCCGTCCAA---
     C13 ACGGT--------CGAAGGA AA----AAGTCTTTCTCGCG AGAGAGAGG----------- -----------------CTG GCTCCTTTTCCGTCCAA---
      S2 ACGGA--------TAGTAGT AA----AAAAA-------- -------------------- ------GTCTTC-------G TCT-GTCAAT-----A----
      S4 ACGGA--------GGACCGT AA----AAAA---------- -------------------- ------GTCTTC------TG TCT-----A-----------
     AGL ACGGATA-----------GT AGC----CCTCTCGGGGGC- -------------------- ------GAGGTCC------G TCTGCTATAC----------
     AGI ACGGATA-----------GT AGC----CCTCTCGGGGGC- -------------------- ------GAGGTCC------G TCTGCTATAC----------
     AGC ACGGATA-----------GT AGC----CCTTTCGGGGGC- -------------------- ------GAGGTCC------G TCTGCTATAC----------
     AGA ACGGATA-----------GT AGT----CCTTTCGGGGGC- -------------------- ------GAGGTCC------G TCTGCTATAC----------
    CAG2 ACGGATA-----------GT AGT----CCTTTCGGGGGC- -------------------- ------GAGGTCC------G TCTGCTATAC----------
     AGD ACGGATA-----------GT AGC----CCTTTCGGGGGC- -------------------- ------GAGGTCC------G TCTGCTATAC----------
     AGE ACGGATA-----------GT AGC----CCTTTCGGCCGC- -------------------- ------GAGGTCC------G TCTGCTATAC----------
     AGR ACTG---------TGAGGGC AT------------------ -------------------- ----GCAAGTGCC------- -CTTATTTAC----------
     AGS ACTG---------TGAGGGC AT------------------ -------------------- ----GCAAGTGCC------- -CTTATTTAC----------
THARZIANU CCAGGCGCCCGCCGGAGGGA CCAACCAAA----------- -------------------- --------CT--C------- TTTTCTGTAGTCCCC-----
```

[illegible]5	--TATA-AACCCAATTTATC	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATC-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
[illegible]39	--TATA-AACCCAATTTATT	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATC-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
[illegible]207	--TATA-AACCCAATTTATT	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATG-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
NO	--TATA-AACCCAATTTATC	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATC-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
[illegible]306	--TATA-AACCCAATTTATT	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATC-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
GU1	--TATA-AACCCAATTTATC	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATC-	-TAAT--ACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
AJ1101	--TATA-AACCCAATTTATT	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATC-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
G4	--TATA-AACCCAATTTATT	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATC-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
RGR39	--TATA-AACCCAATTTATT	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATG-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
L91S	--TATA-AACCCAATTTATT	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATC-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
RGR38	--TATA-AACCCAATTTATT	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATG-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
SLP33	--TATA-AACCCAATTTATT	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATC-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
9215512	--TATA-AACCCAATTTATT	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CCCATC-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
K1	--TATA-AACCCAATTTATT	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATC-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
9215513	--TATA-AACCCAATTTATT	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CCCATC-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
Sdos	--TATA-AACCCAATTTATT	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATC-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
R64	--TACA-AACCCAATTTATT	TTAAAAC-GAATGTAA-TGG	------ATGTAA-CGCATC-	-TAAT--ACTAAGTTTCAAC	AACGCATCTCTTGGCTCTCG
R56	--TATA-AACCCAATTTATT	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATC-	-TAAT--ACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
R65	--TATA-AACCCAATTTATT	TTAAAAT-GAATGTAA-TGG	------ATGTAA-CGCATC-	-TAAT--ACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
S5	--CACA-AACCCCATTGTAT	TTAAATT-GAATGTAA-TTG	------ATGTAA-CGCATCA	-TTAG-AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
S6	--CACA-AACCCCATCGTAT	TTAAGTT-GAATGTAA-TTG	------ATGTAA-CGCATCA	-TTAG-AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
F1	--CACA-AACCCCATTGTAT	TTAAATT-GAATGTAA-TTG	------ATGTAA-CGCATCA	-TTAA-AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
F2	--CACA-AACCCCATTGTAT	TTAAATT-GAATGTAA-TTG	------ATGTAA-CGCATCA	-TTAGGAACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
F3	--CACA-AACCCCATTGTAT	TTAAATT-GAATGTAA-TTG	------ATGTAA-CGCATCA	-TTAG-AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
F5	--CACA-AACCCCATTGTAT	TTAAATT-GAATGTAA-TTG	------ATGTAA-CGCATCA	-TTAG-AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
F7	--CACA-AACCCCATTGTAT	TTAAATT-GAATGTAA-TTG	------ATGTAA-CGCATCA	-TTAG-AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
F6	--CACA-AACCCCATTGTAT	TTAAATT-GAATGTAA-TTG	------ATGTAA-CGCATCA	-TTAG-AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
S1	--CACA-AACCCCATTGTAT	TTAAATT-GAATGTAA-TTG	------ATGTAA-CGCATCA	-TTAG-AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
A2	--TATA-AACTCAATTTATT	TTAAAAC-GGATGTAA-TGG	------ATGTAA-CACATC-	-TAA--TACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
A1	--TATA-AACTCAATTTATT	TTAAAAC-GGATGTAA-TGG	------ATGTAA-CACATC-	-TAA--TACTAAGTTGCAAC	AACGGATCTCTTGGCTCTCG
B	--TATA-AACTCAATTTATT	TTAAAAC-GGATGTAA-TGG	------ATGTAA-CACATC-	-TAA--TACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
AGK	--TATACAAACCCATT-TAA	TTGAACT-GAATGTAAATTG	------ATGTAA-CGCATCT	TTA---AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
AGN	--TATACAAACTCATT-TAA	TTGAACT-GAATGTAC-TTG	------ATGTAA-CGCATCA	TTA---AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
AGP	--TATACAAACTCATT-TAA	TTGAACT-GAATGTAC-TTG	------ATGTAA-CGCATCA	TTA---AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
AGBO	--TATACAAACTCATT-TAA	TTGAACT-GAATGTAC-TTG	------ATGTAA-CGCATCA	TTA---AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
AGBB	--TATACAAACTCATT-TAA	TTGAACT-GAATGTAC-TTG	------ATGTAA-CGCATCA	TTA---AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
S3	--TATACAAACTCATT-TAA	CTGCACT-GAATGTAC-TTG	------ATGTAA-CGCATCA	TTAC---ACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
AGJ	---ACACAAACTCATT-TAA	TTAAATT-GAATGTAA-TTG	------ATGTAA-CGCATCA	TTAG--AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
AGG	---ACACAAACTCATTGTAT	TTAAACT-GAATGTAA-TTG	------ATGTAA-CGCATCA	TTAG--AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
C5	--CACA-AACTCTATTTATT	TTAAAAC-GAATGTAA-TTG	------ATGTAA-CGCATC-	-TA--ATACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
C8	--CACA-CACTCTATTTATT	TTAAAAT-GAATGTAA-CAG	------ATGTAA-CGCATC-	-TA--TTACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
C2	--CACA-CACTCTATTTAAT	TTGAACT-GAATGTAA-TTG	------ATGTAA-CGCATC-	-TA--ATACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
C4	--CACA-CACTCTATTTAAT	TTAAACT-GAATGTAA-TTG	------ATGTAA-CGCATC-	-TA--ATACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
C6	--TATA-AAATCAATTTAAT	TTAAAAC-GAATGTAA-TGG	------ATGTAA-CGCATC-	-TA--ATACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
C1	--CACA-CACTCTATTTAAT	TTGAACT-GAATGTAA-TTG	------ATGTAA-CGCATC-	-TA--ATACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
C11	TACATA-AAATCTTATATAT	CTAATCA-GAATGTAA-TCG	------ATGTAAACGCATC-	-TA-TAAACTAAGTTTCAAC	AACGGATCTCGTGGCTCTCG
C12	TACATA-AAATCTTATATAT	TTAATCA-GAATGTAA-TCG	------ATGTAAACGCATC-	-TA-TAAACTAAGTTTCAAC	AACGGATCTCGTGGCTCTCG
C13	TACATA-AAATCTTATATAT	CTAATCA-GAATGTAA-TCG	------ATGTAAACGCATC-	-TA-TAAACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
S2	--CATA-AACTCCAATTTAA	TAAATCT-GAATGTAA-TTG	------ATGTAA-CACATCT	TTAA---ACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
S4	--CATA-AACTCCAATTTAA	TAAATCT-GAATGTAA-TTG	------ATGTAA-CACATCT	TTAA---ACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
AGL	---ATA-AACTCCAG--TAT	ATAAATTTGAATGTCATTTG	------ATGTAA-CACATCT	ATA---AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
AGI	---ATA-AACTCCAG--TAT	ATAAATTTGAATGTCATTTG	------ATGTAA-CACATCT	ATA---AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
AGC	---ATA-AACTCCAA--TAT	ATAAATTTGAATGTAATTTG	------ATGTAA-CGCATCT	TTA---AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
AGA	---ATA-AACTCCAA--TAT	ATAAATG-GAATGTAATTTG	------ATGTAA-CGGATCT	TTA---AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
CAG2	---ATA-AACTCCAA--TAT	ATAAATG-GAATGTAATTTG	------ATGTAA-CGGATCT	TTA---AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
AGD	---ATA-AACTCCAA--TAT	ATAAATTTGAATGTAATTTG	------ATGTAA-CGCATCT	TTA---AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTCG
AGE	---ATA-TACTCCAA--TAT	ATAAATTTGAATGTAATTTG	------ATGTAA-CGCATCT	TTA---AACTAAGTTTCAAC	AACGGATCTCTTGGCTCTTG
AGR	--TACATACACCTAA--TAT	GGAAATC-GCACGTGACTTG	CGTCGTATGGCG-CTTAAAA	CTAAA-AAC-AAGTTTCAAC	AACGGATCTCTTGGCTCTCG
AGS	--TACATACACCTAA--TAT	GGAAATC-GCACGTGACTTG	CGTCGTATGGCG-CTTAAAA	CTAAA-AAC-AAGTTTCAAC	AACGGATCTCTTGGCTCTCG
THARZIANU	--TCGCGGACGTTATTTCTT	ACAGCTCTGAGCAAAAATTC	--------AAAA-TGAATC-	--AA------AACTTTCAAC	AACGGATCTCTTGGTTCTGG

5	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
RGR-59	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
9207	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
NO	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
C306	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
GUL	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AJ1101	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
G4	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
RGR39	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
L91S	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
RGR38	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
SLP33	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
9215512	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
K1	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
9215513	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
Sdos	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
R64	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
R56	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
R65	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
S5	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
S6	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
F1	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
F2	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
F3	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
F5	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
F7	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
F6	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
S1	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
A2	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
A1	CATCGATGAAAAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
B	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AGK	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AGN	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AGP	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AGBO	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AGBB	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
S3	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AGJ	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AGG	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
C5	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
C8	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
C2	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
C4	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
C6	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
C1	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
C11	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
C12	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
C13	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
S2	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
S4	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AGL	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AGI	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AGC	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AGA	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
CAG2	CATCGATGAAGAACGCAGCG	AAACCCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AGD	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AGE	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AGR	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
AGS	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACC	TTGCGCTCCTTGGTATTCCT
THARZIANU	CATCGATGAAGAACGCAGCG	AAATGCGATAAGTAATGTGA	ATTGCAGAATTCAGTGAATC	ATCGAATCTTTGAACGCACA	TTGCGCCCGCCAGTATTCTG

5	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAACCTTTTTGTTAACTCAA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
Gn-59	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAACCTTTTTGTTAACTCAA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
9207	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAACCTTTTTGTTAACTCAA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
NO	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAACCTTTTTGTTAACTCAA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
C306	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAACCTTTTTGTTAACTCAA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
GU1	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAACCTTTTTGTTAACTCAA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
AJ1101	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAATCTTTTTGTTAACTCGA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
G4	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAATCTTTTTGTTAACTCGA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
RGR39	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAATCTTTTTGTTAACTCGA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
L91S	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAATCTTTTTGTTAACTCGA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
RGR38	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAATCTTTTTGTTAACTCGA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
SLP33	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAATCTTTTTGTTAACTCAA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
9215512	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAATCTTTTTGTTAACTCAA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
K1	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAATCTTTTTGTTAACTCAA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
9215513	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAATCTTTTTGTTAACTCAA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
Sdos	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAATCTTTTTGTTAACTCAA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
R64	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAATCTTTTTGTTAACTCAA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
R56	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAATCTTTTTGTTAACTCAA	TTTGGTTT---CACTTT-GG	TATTGGAGGT-TCTTGCAGC
R65	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	-AAATCTTTTGTTAATTCAA	CTGGT-T----TGCTTT-GG	ATTTGGAGG--TTTTGCAGA
S5	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	-AAATCTTTTGTTAATTCAA	CTGGT-T----TGCTTT-GG	ATTTGGAGG--TTTTGCAGA
S6	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	-AAATTTTTTGTTAATTCAA	CTGGT-T----TGCTTT-GG	ATTTGGAGG--TTTTGCAAA
F1	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	-AAATCTTTTGTTAATTCAA	CTGGT-T----TGCTTT-GG	ATTTGGAGG--TTTTGCAGA
F2	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	-AAATCTTTTGTTAATTCAA	CTGGT-T----TGCTTT-GG	ATTTGGAGG--TTTTGCAGA
F3	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	-AAATCTTTTGTTAATTCAA	CTGGT-T----TGCTTT-GG	ATTTGGAGG--TTTTGCAAA
F5	TGGAGCATGCCTGTTTGAGT	ATCATGAAATTTCTCAAAGT	-AAATCTTTTGTTAATTCAA	CTGGT-T----TGCTTT-GG	ATTTGGAGG--TTTTGCAAA
F7	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	-AAATTTTTTGTTAATTCAA	CTGGT-T----TGCTTT-GG	ATTTGGAGG--TTTTGCAAA
F6	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	-AAATCTTTTGTTAATTCAA	CTGGT-T----TGCTTT-GG	ATTTGGAGG--TTTTGCAGA
S1	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	-AAACCTTTTGTTAATTCAA	CTGGTCTT---TGCTTT-GG	ATTTGGAGGCCT-TTGCAGA
A2	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	CAAACCTTTTGTTAATTCAA	TTGGT-TC---TGCTTT-GG	TATTGGAGGATTATTGCAGC
A1	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	CAAACCTTTTGTTAATTCAA	TTGGT-TC---TGCTTT-GG	TATTGGAGGATTATTGCAGC
B	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	AAAC-CTTTTGTTAATTCAA	CTGGTCT---TTGCTTT-GG	ATTTGGAGGCCT-TTGCAGA
AGK	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	AAAC-CTTTTGTTAATTCAA	TTGGCCTT----GCTTT-GG	ATTTGGAGGTCTT--GCAGA
AGN	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	AAAC-CTTTTGTTAATTCAA	CTGGCCTT----GCTTT-GG	ATTTGGAGGTCTT--GCAGA
AGP	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	AAAC-CTTTTGTTAATTCAA	TTGGCCTT----GCTTT-GG	ATTTGGAGGTCTT--GCAGA
AGBO	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	AAAC-CTTTTGTTAATTCAA	TTGGCCTT----GCTTT-GG	ATTTGGAGGTCTT--GCAGA
AGBB	TGGAGCATGCCTGTTTGAGT	ATCATGAATTT-CTCAAAGT	AAAC-CTTTTGTTAATTCAA	TTGAT-T----TTATTTTGG	ACTTGGGGG--TCTTGCAGA
S3	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	-AAACCTTTTGTTCATTCGA	CTGGGTTT----GCTTT-GG	ACTTGGAGGTCT-TTGCAGA
AGJ	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	AAAT-CTTTTGTTAATTCAA	CTGGTTTT----GCTTT-GG	ACTTGGAGGTCT-TTGCAGA
AGG	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	AAAT-CTTTTGTTAATTCAA	TTGGT-TCC-TTACTTT-GG	TATTGGAGGTATATTGCAGC
C5	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	-AAACCTTTTGTTAATTCAA	TTGGTTTCT---GCTTT-GG	TATTGGAGGTATATTGCAGC
C8	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	-CAATCTTTTGTTAATTCAA	TTGGT-TCC-TCACTTT-GG	TTTTGGAGGTCT-TTGCAGT
C2	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	-AAACCTTTTGTTAATTCAA	TTGGT-TCC-TTGCTTT-GG	TATTGGAGGTATATTGCAGC
C4	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	-CAATCTTTTGTTAACTCAA	TTGGTTTCT---ACTTT-GG	TATTGGAGGT-TATTGCAGC
C6	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-TTCAAAGT	-AAATCTTTTGTTAATTCAA	TTGGC-TCC-TTGTTTT-GG	TTTTGGAGGTCT-TTGCAGC
C1	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGT	-ATACCTTTTGTTAATTCAA	CTGGCTT---TTGCTTT-GG	ATTTGGAGGTCT--TGCAGA
C11	CGGAGCACGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGC	-AAGTCTTTTGTTAATTCAA	CTGGCTT---TTGCTTT-GG	ATTTGGAGGTCT--TGCAGA
C12	CGGAGCACGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGC	-AAGTCTTTTGTTGATTCAA	CTGGCTT---TTGCTTT-GG	ATTTGGAGGTCT--TGCAGA
C13	CGGAGCACGCCTGTTTGAGT	ATCATGAAATT-CTCAAAGC	-AAGTCTTTTGTTAATTCAA	CTGGCTT---TTGTTTT-GG	ATTTGGAGGTCT--TGCAGA
S2	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-ATCAAAGT	-AAATCTTTTGTTCATTCGA	TTGAT-T----TTATTTTGG	ACTTGGGGG--TCT-GCAGA
S4	TGGAGCATGCCTGTTTGAGT	ATCATGAAATT-ATCAAAGT	-AAATCTTTTGTTCATTCGA	TTGAT-T----TTATTTTGG	ACTTGGGGG--TCT-GCAGA
AGL	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-CTCAAAAT	AAAT-CTTTTGTTCATTCGA	TCGATTTT-----ATTTTGG	ACTTGGAGGTCT---GCAGA
AGI	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-CTCAAAAT	AAAT-CTTTTGTTCATTCGA	TTGATTTT-----ATTTTGG	ACTTGGAGGTCT---GCAGA
AGC	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-CTCAAAAT	AAAT-CTTTTGTTCATTCGA	TTGATTTT----ATTTT-GG	ACTTGGAGGTCT---GCAGA
AGA	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-CTCAAAAT	AAAT-CTTTTGTTCATTCGA	TTGATTTT-----ATTTTGG	ACTTGGAGGTCT---GCAGA
CAG2	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-CTCAAAAT	AAAT-CTTTTGTTCATTCGA	TTGATTTT-----ATTTTGG	ACTTGGAGGTCT---GCAGA
AGD	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-CTCAAAAT	AAAT-CTTTTGTTCATTCGA	TTGATTTT-----ATTTTGG	ACTTGGAGGTCT---GCAGA
AGE	TGGAGCATGCCTGTTTGAGT	ATCATGAAATC-CTCAAAAT	AAAT-CTTTTGTTCATTCGA	TTGATTTT-----ATTTTGG	ACTTGGAGGTCT---GCAGA
AGR	TGGAGCACGCCTGTTTGAGT	ATCATGAAACT-CTCAAGCT	AGATGCTTTAGTGTCATCTA	CTTGGTTA-------TT-GG	ACTCTGCTGTCT-----CCT
AGS	TGGAGCACGCCTGTTTGAGT	ATCATGAAACT-CTCAAGCT	AGATGCTTTAGTGTCATCTA	CTTGGTTA-------TT-GG	ACTCTGCTGTCT-----CCT
THARZIANU	GCGGGCATGCCTGTCCGAGC	GTCATTTCAACCCTCGAACC	CCTCCGGGGGGTCGGCGTTG	GGGACCTCGGGAGCCCCTAA	GACGGGATCCCGGCCCCGAA

```
C116      TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCT--
RGR-59    TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCC--
9207      TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCT--
NO        TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCT--
C306      TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCC--
GU1       TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCT--
AJ1101    TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CACTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCC--
G4        TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCC--
RGR39     TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCC--
L91S      TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCC--
RGR38     TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCC--
SLP33     TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCC--
9215512   TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCCC-
K1        TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCCC-
9215513   TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CA-TGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCCC-
Sdos      TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCCC-
R64       TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCTGA-GGACACCCC-
R56       TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCCGA-GGACACCCC-
R65       TT-CACACGC-TGCTCCTCT TTGTTCA-TTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAATTATC TATCGCCGA-GGACACCCC-
S5        TTTCACAGTC-TGCTCCTCT TAAATAAATTAGCTGGATCT CAGTAT-ATGCTTGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACAC--T-
S6        TTTCACAGTC-TGCTCCTCT TAAATAAATTAGCTGGATCT CAGTAT-ATGCTTGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACAC--T-
F1        TTTCACAGTC-TGCTCCTCT TAAATAAATTAGCTGGATCT CAGTAT-ATGCTTGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGAAGGACAC--T-
F2        TTTCACAGTC-TGCTCCTCT TAAATAAATTAGCTGGATCT CAGTAT-ATGCTTGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACAC--T-
F3        TTTCACAGTC-TGCTCCTCT TAAATAAATTAGCTGGATCT CAGTAT-ATGCTTGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACAC--T-
F5        TTTCACAGTC-TGCTCCTCT TAAATAAATTAGCTGGATCT CAGTAT-ATGCTTGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACAC--T-
F7        TTTCACAGTC-TGCTCCTCT TAAATAAATTAGCTGGATCT CAGTAT-ATGCTTGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACAC--T-
F6        TTTCACAGTC-TGCTCCTCT TAAATAAATTAGCTGGATCT CAGTAT-ATGCTTGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACAC--T-
S1        TTTCACAGTC-TGCTCCTCT TAAATAAATTAGCTGGATCT CAGTAT-ATGCTTGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACAC--T-
A2        TGTCACAGTC-TGCTCCTCT TAAATAAATTAGCTGGATCT CAGTAT-ACACTTGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACAC--T-
A1        T-TCACAC-C-TGCTCCTCT TTG-TGCATTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCAGCGTGATAAGTTATC TATCGCTGA-GGACACCCT-
B         T-TCAC-ACC-TGCTCCTCT TTG-TGCATTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCAGCGTGATAAGTTATC TATCGCTGA-GGACACCCT-
AGK       TGTCACAGTC-TGCTCCTCT TAAATAAATTAGCTGGATCT CAGTATAC-ACTTGGTTC-C ACTCGGCGTGATAAG-TATC ACTCGCTGA-GGACAC--T-
AGN       TGTCACAGTC-TGCTCCTCT TAAATGAATTAGCTGGATCT CAGTATAC-ACTTGGTTC-C ACTCGGCGTGATAAG-TATC ACTCGCTGA-GGACACTGT-
AGP       TGTCACAGTC-TGCTCCTCT TAAATGAATTAGCTGGATCT CAGTATAC-ACTTGGTTC-C ACTCGGCGTGATAAG-TATC ACTCGCTGA-GGACACTGT-
AGBO      TGTCACAGTC-TGCTCCTCT TAAATATATTAGCTGGATCT CAGTATAC-ACTTGGTTC-C ACTCGGCGTGATAAG-TATC ACTCGCTGA-GGACACTGT-
AGBB      TGTCACAGTC-TGCTCCTCT TAAATATATTAGCTGGATCT CAGTATAC-ACTTGGTTC-C ACTCGGCGTGATAAG-TATC ACTCGCTGA-GGACACTGT-
S3        TGTCACAGTC-TGCTCCTCT TAAATATATTAGCTGGATCT CAGTAT-ACACTTGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACAC--T-
AGJ       TTTCAC-GTC-TGCTCCTCT TAAATGCATTAGCTGGATCT CAGTATAT-GCTTGGTTC-C ACTCGGCGTGATAAG-TATC ACTCGCTGA-GGACACCTT-
AGG       TTTCAC-GTC-TGCTCCTCT TAAATGCATTAGCTGGATCT CAGTATAT-GCTTGGTTC-C ACTCGGCGTGATAAG-TATC ACTCGCTGA-GGCCACTGT-
C5        TTTCACACCC-TGCTCCTCT TT-GTTCATTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAGTTATC TATCGCTGA-GGACAC--T-
C8        TTTCACACCC-TGCTCCTCT TT-GTACATTAGCTGGATCT CGGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAGTTATC TATCGCTGA-GGACAC--T-
C2        TT-CACAAC--TGCTCCTCT TT-GTTCATTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAGT-ATC TATCGCTGA-GGACAC--T-
C4        TTTCACACCC-TGCTCCTCT TT-GTTCATTAGCTGGATCT CGGTGTTATGCTTGGTTC-C ACTCGGCGTGATAAGTTATC TATCGCTGAAGGACAC--T-
C6        TT-CACACC--TGCTCCTCT TT-GTGCATTAGCTGGATCT CAGTGTTATGCTTGGTTC-C ACTCAGCGTGATAAGTTATC TATCGCTGA-GGACAC----
C1        TT-CAC-GTC-TGCTCCTCT TTTGTTCATTAGCTGGATCT CAGTATAAAACTCGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACACTCT-
C11       TT-CAC-GTC-TGCTCCTCT TAAATGCATTAGCTGGATCT C--TATAAAAC-CGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACACTCT-
C12       TT-CAC-GTC-TGCTCCTCT TAAATGCATTAGCTGGATCT C--TATAAAAC-CGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACACTCT-
C13       TT-CAC-GTC-TGCTCCTCT TAAATGCATTAGCTGGATCT C--TATAAAAC-CGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACACTCT-
S2        T-TCAC-GTC-TGCTCCCCT TAAATTCATTAGCTGGATCA CTGTAG-A--CTCGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACAC--T-
S4        T-TCAC-GTC-TGCTCCCCT TAAATTCATTAGCTGGATCA CTGTAG-A--CTCGGTTC-C ACTCGGCGTGATAAGT-ATC ACTCGCTGA-GGACAC--T-
AGL       T-TCAC-GTC-TGCTCCTCT TAAATTCATTAGCTGGATCT GTATGAA---CTCGGTTC-C ACTCGGCGTGATAAG-TATC ACTCGCTGA-GGACACTGT-
AGI       T-TCAC-GTC-TGCTCCTCT TAAATTCATTAGCTGGATCT GTATGAA---CTCGGTTC-C ACTCGGCGTGATAAG-TATC ACTCGCTGA-GGACACTGT-
AGC       T-TCAC-GTC-TGCTCCTCT TAAATGTATTAGCTGGATCT CAGTGAA---CTCGGTTC-C ACTCGGCGTGATAAG-TATC ACTCGCTGA-GGACACTGT-
AGA       T-TCAC-GTC-TGCTCCTCT TAAATGTATTAGCTGGATCT CAGTGAA---CTCGGTTC-C ACTCGGCGTGATAAG-TATC ACTCGCTGA-GGACACTGT-
CAG2      T-TCAC-GTC-TGCTCCTCT TAAATGTATTGGCTGGATCT CAGTGAA---CTCGGTTC-C ACTCGGCGTGATAAG-TATC ACTCGCTGA-GGACACTGT-
AGD       T-TCAC-GTC-TGCTCCTCT TAAATGTATTAGCTGGATCT CAGTGAA---CTCGGTTC-C ACTCGGCGTGATAAG-TATC ACTCGCTGA-GGACACTGT-
AGE       T-TCAC-GTC-TGCTCCTCT TAAATGTATTAGCTGGATCT CAGTGAA---CTCGGTTC-C ACTCGGCGTGATAAG-TATC ACTCGCTGA-GGACACTGT-
AGR       TGTCTGGAAC-GGCTGGTCT CAAATGTATTAGCTGGTCCT TGTTGTGGGATTTGGTTC-T ACTCAGCGTGATAATCTATG AC-CGCTGA-GGACATCTT-
AGS       TGTCTGGAAC-GGCTGGTCT CATATGTATTAGCTGGTCCT TGTTGTGGGATTTGGTTC-T ACTCAGCGTGATAATCTATG AC-CGCTGA-GGACATCTT-
THARZIANU ATACAGTGGCGGTCTCGCCG CA---GCCTCTCCTG---CG CAGTAGTTTGCACAACTCGC ACCGGGAGCGCGGCGC-GTC CACGTCCGTAAAACACCCAA
```

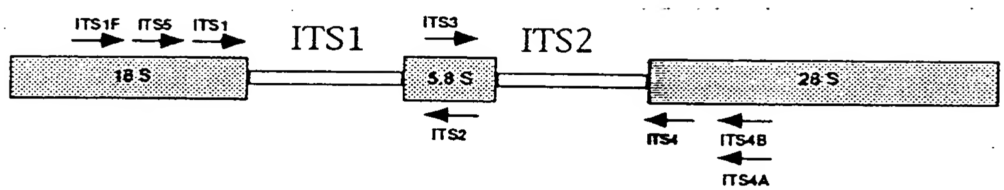
6	GATAAAAAG-GG--TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTCCTAATA-GTCC	ATTGGACTTGG-ACAACAA-	--TAAAATA-ATAATTATTT
[illegible]59	GATGAAAAG-GT--TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTCCTAATA-GTCC	ATTGGACTTGG-ACAACAA-	--TAAAATA-ATAATTATTT
[illegible]9207	GATAAAAAG-GT--TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTCCTAATA-GTCC	ATTG-ACTTGG-ACAA----	--TAAAATA-ATCATTATTT
NO	GATAAAAAG-GG--TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTG-ACTTGG-ACAA----	--TAAAATA-ATTATTATTT
C306	GATAAAAAG-GT--TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTG-ACGTGG-ACAA----	--TAAAATA-ATAATTATTT
GU1	GATAAAAAG-GT--TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTCCTAATA-GTCC	ATTG-ACTTGG-ACAACAA-	--TAAAATA-ATAATTATTT
AJ1101	GGTAAAAAA-GG--TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTC-ACTTGG-ACAA----	--TAAAAT----TATTATTT
G4	GGTAAAAAA-GG--TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTC-ACTTGG-ACAA----	--TAAAAT----TATTATTT
RGR39	GGTAAAAAA-GG--TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTC-ACTTGG-ACAA----	--TAAAAT----TATTATTT
L91S	GGTAAAAAA-GG--TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTC-ACTTGGGACAA----	--TAAAAT----TATTATTT
RGR38	GGTAAAAAA-GG--TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTC-ACTTGG-ACAA----	--TAAAAT----TATTATTT
SLP33	GGTAAAAAA-GG--TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTC-ACTTGGGACAA----	--TAAAAT----TATTATTT
9215512	GGTAAAAAA-GGG-TGGCCA	AGGTAAATAC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATGG-ACTTGG-ACAA----	--TAAAAA---TTATTATTT
K1	GGTAAAAAA-GGG-TGGCCA	AGGTAAATAC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTG-ACTTGG-ACAA----	--TAAAAA---TTATTATTT
9215513	GGTAAAAAA-GGG-TGGCCA	AGGTAAATAC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATGG-ACTTGG-ACAT----	--AAAAAT----TATTATTT
Sdos	GGTAAAAAA-GG--TGGCCA	AGGTAAATAC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTG-ACTTGG-ACAA----	--TAAAAA---TTATTATTT
R64	GGTAAAAAA-GGGGTGGCCA	AGGTAAATGC---AAAT-GA	--ACCGCTTCTAATA-GTCC	ATTG-ACTTGG-ACAA----	--TAAAAA---TTATTATTT
R56	GGTAAAAAA-GGG-TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTG-ACTTGG-ACAA----	--TAAAAAA--TTATTATTT
R65	GGTAAAAAA-GGG-TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTG-ACTTGG-ACAA----	--TAAAAAA--TTATTATTT
S5	-GTAAAA----GG-TGGCCA	--GGAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTA-AATTGG-ACAAA---	--TTAAA---TTTATGATCT
S6	-GTAAAA---GG--TGGCCA	--GGAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTA-AATTGG-ACAAA---	--TTAAA---TTTATGATCT
F1	-GTAAAA---GG--TGGCCA	--GGAAATGC---AAAT-AA	--ACCGCTTCTAATA-GTCC	ATTA-AATTGG-ACAAA---	--TTAAA---TTTATGATCT
F2	-GTAAAA---GG--TGGCCA	--GGAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTA-AATTGG-ACAAA---	--TTAAA---TTTATGATGT
F3	-GTAAAA---GG--TGGCCA	--GGAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTA-AATTGG-ACAAA---	--TTAAA---TTTATGATCT
F5	-GTAAAA---GG--TGGCCA	--GGAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTA-AATTGG-ACAAA---	--TTAAA---TTTATGATCT
F7	-GTAAAA---GG--TGGCCA	--GGAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTA-AATTGG-ACAAA---	--TTAAA---TTTATCATCT
F6	-GTAAAA---GG--TGGCCA	--GGAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTA-AATTGG-ACAAA---	--TTCAA---TTTATGATCT
S1	-GTAAAA---GG--TGGCCA	--GGAAATGC---AGAT-GA	--ACCGCTTCTAATAAGTCC	ATTA-AATTGG-ACAAA---	--TTAAA---TTTATGATCT
A2	GTTAAAAA--GG--TGGCCA	GG--CAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTGGACTTGG-ACAAAT--	--TAAAAA--TTTATGATCT
A1	GTTAAAAAGGGG--TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTGGACTTGG-ACAATA--	--TCTAA---TTTATGATCT
B	GTTAAAAAGGGG--TGGCCA	AGGTAAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTG-ACTTGG-ACAATA--	--TCTAA---TTTATGATCT
AGK	GTTAAAAA--GG--TGGCCA	--GGCAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTG-ACTTGG-ACAAA---	--TTAAAAA-TTTATGATCT
AGN	---AAAA---GG--TGGCCA	--GGCAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTC-ACTTGG-ACAATT--	----ACA---TTTATGATCT
AGP	---AAAA---GG--TGGCCA	--GGCAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTC-ACTTGG-ACAATT--	----ACA---TTTATGATCT
AGBO	---AAAA---GG--TGGCCA	--GGCAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTC-ACTTGG-ACAATT--	----ACA---TTTATGATCT
AGBB	---AAAA---GG--TGGCCA	--GGCAATGC---AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTC-ACTTGG-ACAA----	--TTACA---TTTATGATCT
S3	-GTAAAAAA-GG--TGGCCA	--GGCAATACT--AGAT-GA	--ACCGCTTCTAATA-GTCC	ATTC-ACTTGG-ACAA----	--TTACA---TTTATGATCT
AGJ	---AAAAA---G--TGGCCA	--GGAAATAC---AGAT-GA	--ACCGCTTCTAATA-GTCT	ATTA-AGTTAG-ACAATT--	-----AA---TTTAAGATCT
AGG	---AAAA---GG--TGGCCA	--GGAAATAC---AGAT-GA	C-ACCGCTTCTAATA-GTCT	ATTA-AGTTAG-ACAATTT-	-----AAT--TTTAAGATCT
C5	TGTAAAAT--GG--TGGCCA	AGGTAAACGCA---GAT-GA	--ACCGCTTCTAACA-GTCC	ATTG-ACTTGG-ACAATA--	---AATA---TTTATGATCT
C8	GTTAAAAAT-GG--TGGCCA	AGGTAAACACA--TGATTGA	--ACCGCTTCTAACG-GTCC	ATTG-ACTTGG-ACAAGA--	--TTAAATT-TTTATGATCT
C2	TGTAAAAA--GG--TGGCCA	AGGTAAACGCA---GAT-GA	--ACCGCTTCTAACCAGTCC	ATTG-ACTTGG-ACAATA--	----AAATA-TTTATGATCT
C4	TGTAAAAAT-GG--TGGCCA	AGGTAAACGCA--TGATTGA	--ACCGCTTCCAACG-GTCC	ATTG-ACTTGG-ACAATA--	-------CTCTTTATGATTT
C6	TGTAAAA---GG--TGGCCA	AGGTAAATGCA---GAT-GA	--ACCGCTTCTAATA-GTCC	ATTG-ACTTGG-ACAAGA--	--TTAAATT-TTTATGATCT
C1	TGTAAAAG--GG--TGGCCA	AGATTCATGCA---GAT-GA	--ACCGCTTCTAACG-GTCT	ATTA-AATTAG-ACAAAA--	--TCACATC-TTTATGATCT
C11	TGCAAAAG--GG--TGGCCG	-GATTCATGGA------CGA	--ACCGCTTCTAACG-GTCT	ATTA-GATTAG-ACAAA---	---CACA--CTTTATGATCT
C12	TGCAAAAG--GG--TGGCCG	-GATTCATGGA------CGA	--ACCGCTTCCAACG-GTCT	ATTA-GATTAG-ACAAA---	---CACA--CTTTATGATGT
C13	TGCAAAAG--GG--TGGCCG	-GATTCATGGA------TGA	--ACCGCTTCTAACG-GTCT	ATTA-GATTAG-ACAAA---	---CACA--CTTTATGATCT
S2	CGTAAAAAA-AG--TGGCC-	--GGGATTACT--AGA--GA	--ACTGCTTCTAATA-GTCC	ATTA-GCTTGG-ACAA----	--CTACA---TTTATGATCT
S4	CGTAAAAAA-AG--TGGCC-	--GGGATTACT--AGA--GA	--ACTGCTTCTAATA-GTCC	ATTA-GCTTGG-ACAA----	--CTACA---TTTATGATCT
AGL	---AAAA---GG--TGGCC-	--GGGATTATT----AT-GA	--ACCGCTTCTAATA-GTCC	ATTG-ACTTGG-ACAAAC--	---TAC-----TTATGATCT
AGI	---AAAA---GG--TGGCC-	--GGGATTGAT---TAT-GA	--ACCGCTTCTAATA-GTCC	ATTG-ACTTGG-ACAAAC--	---TAC-----TTATGATCT
AGC	---AAAA---GG--TGGCC-	--GGGGTTGC---AGAC-GA	--ACCGCTTCTAATC-GTCC	ATTG-ACTTGG-ACAATT--	---TA-----TTTATGATCT
AGA	---AAAA---GG--TGGCC-	--GGGGTTGC---AGAC-GA	--ACCGCTTCTAATC-GTCC	ATTG-ACTTGG-ACAATT--	---TA-----TTTATGATCT
CAG2	---AAAA---GG--TGGCC-	--GGGGTTGC---AGAC-GA	--ACCGCTTCTAATC-GTCC	ATTG-ACTTGG-ACAATA--	---TA-----TTTATGATCT
AGD	---AAAA---GG--TGGCC-	--GGGGTTGC---AGAC-GA	--ACCGCTTCTAATC-GTCC	ATTG-ACTTGG-ACAATT--	---TA-----TTTATGATCT
AGE	---AAAA---GG--TGGCC-	--GGGGTTGC---AGAC-GA	--ACCGCTTCTAATC-GTCC	ATTG-ACTTGG-ACAGGT--	---TA-----TTTATGATCT
AGR	-------TCGGGATGGCCA	--AGCC-TGCCT----TGGG	---CTGCTTCTAA----CTC	ATCATCAT--G-ATATTT--	----------TTCAACTTT
AGS	-------TCGGGATGGCCA	--AGCC-CGCCT----TGGG	---CTGCTTCTAA----CTC	ATCATCAT--G-ATATTT--	----------TTCAACAAT
THARZIANU	CTTCTGAAATGT--TGACCT	CGGATCAGGT-----AGGAA	TACCCGCTGAACTTA--AGC	ATATCAATAAGCG------	-------------------

```
[illegible]6 TATG---------------- -------------------- -----------------
RGR-59    TATG---------------- -------------------- -----------------
9207      TATG---------------- -------------------- -----------------
NO        TATG---------------- -------------------- -----------------
C306      TATG---------------- -------------------- -----------------
GUL       TATG---------------- -------------------- -----------------
AJ1101    TATG---------------- -------------------- -----------------
G4        TATG---------------- -------------------- -----------------
RGR39     TATG---------------- -------------------- -----------------
L91S      TATG---------------- -------------------- -----------------
RGR38     TATG---------------- -------------------- -----------------
SLP33     TATG---------------- -------------------- -----------------
9215512   TATG---------------- -------------------- -----------------
K1        TATG---------------- -------------------- -----------------
9215513   TATG---------------- -------------------- -----------------
Sdos      TATG---------------- -------------------- -----------------
R64       TATG---------------- -------------------- -----------------
R56       TATG---------------- -------------------- -----------------
R65       TATG---------------- -------------------- -----------------
S5        GATCTCAAATCAGGTTAGGA CTACCCGCTGAATTACT--- -----------------
S6        GATCTCAAACCAGGTTAGGA CTACCCGCTGAATTACT--- -----------------
F1        GATCTCAAATCAGGT-TGGG CTACCCGCTGAACTTAAGCC ATA--------------
F2        GCTCTCAAATCAGGT-AGGA CTACCCGCT----------- -----------------
F3        GATCTCAAATCAGGT-AGGA CTACCCGCT----------- -----------------
F5        GATCTCAAATCAGTCGAGGA CTACCCGCT----------- -----------------
F7        GATCTCAAATCAGGT-AGGA CT------------------ -----------------
F6        GATCTCAAATCAGGT-AGAT -------------------- -----------------
S1        GATCTCAAATCAGGT-AGGA CTACCCGCTGAACTTAAGCA TATCAACA---------
A2        GATCTCAAATCAGGT-AGGA CTACCCGCTGAACTTAAGCA TATCAATAAGCGGA---
A1        GATCTCAAATCAGGT-AGGA CTACCCGCTGAACTTAAGCA TATCAATAAGCGGA---
B         GATCTCAAATCAGGT-AG-- ATA--CGCCTGACTAA---- -----------------
AGK       GATCTCAAATCAGGT-AGGA CTACCCGCTGAACTTAAGCA TATCAAT----------
AGN       GATCTCAAATCAGGT-AGGA CTACCCGCTGAACTTAAGCA TATCAATAAGCCGA---
AGP       GATCTCAAATCAGGT-AGGA CTACCCGCTGAACTTAAGCA TATCAATAAGCGGAGGA
AGBO      GATCTCAAATCAGGT-AGG- -------------------- -----------------
AGBB      GATCTCAAATCAGGT-AGG- -------------------- -----------------
S3        GATCTCAAACCAGGT-ACGA CTACCC-------------- -----------------
AGJ       GATCTCAAATCAGGT-AGGA CTACC--------------- -----------------
AGG       GATCTCAAATCAGGT-AGG- -------------------- -----------------
C5        GATCTCAAATCAGGT-AGGA CTACCCGCTGAACTTACCT- -----------------
C8        GATCTCAAATCAGGT-AGGA CTACCCGCTGAACTTACTAG -----------------
C2        GATCTCAAATCAGGT-AGA- -TACCG-------------- -----------------
C4        GATCTCAAATCAGGT-AGGA CT------------------ -----------------
C6        GATCTCAAATCAGGT-AGGA CTACCCGCTGAACTACT--- -----------------
C1        GATCTCAAATCAGGT-GGGA CTACCCGCTGAAC------- -----------------
C11       GATCTCAAATCAGGT-GGGA CTACCCGCTGAACTTAAAGC ATATCA-TAG-------
C12       GATCTCAAATCAGGT-GGGA CTACCCGCTGAACTTAAAGC ATATCA-TAG-------
C13       GATCTCAAATCAGGT-GGGA CTACCCGCTGAACTTAAAGC ATATCAATAG-------
S2        GATCTCAAATCAGGT-AGGA CTACCC-------------- -----------------
S4        GATCTCAAATCAGGT-AGGA CTACCC-------------- -----------------
AGL       GATCTCAAATCAGGT-AGGA CTACCCGCTGAACTTAAGCA TATCAATAAGCGGAG--
AGI       GATCTCAAATCAGGT-AGGA CTACCCG------------- -----------------
AGC       GATCTCAAATCAGGT-AGG- -------------------- -----------------
AGA       GATCTCAAATCAG------- -------------------- -----------------
CAG2      GATCTCAAATCAG------- -------------------- -----------------
AGD       GATCTCAAATCAGGT-AGG- -------------------- -----------------
AGE       GATCTCAAATCAGGT-AGG- -------------------- -----------------
AGR       GACCTCGAATCAGGC-GGGA CTACCCGCTGAACTTAAGCA TATCAATAAGCGGAGGA
AGS       GACCTCGAATCAGGC-GGGA CTACCCGCTGCCCTTAAGCA TATCAATAAGCGGAGGA
THARZIANU -------------------- -------------------- -----------------

;
end;
```

FIGURE 1

Scheme representing the Internal Transcriber spaces (ITS) region of the ribosomal DNA. The shadowed boxes represent the ribosomal genes. The arrows denote the position of the PCR and sequencing primers.



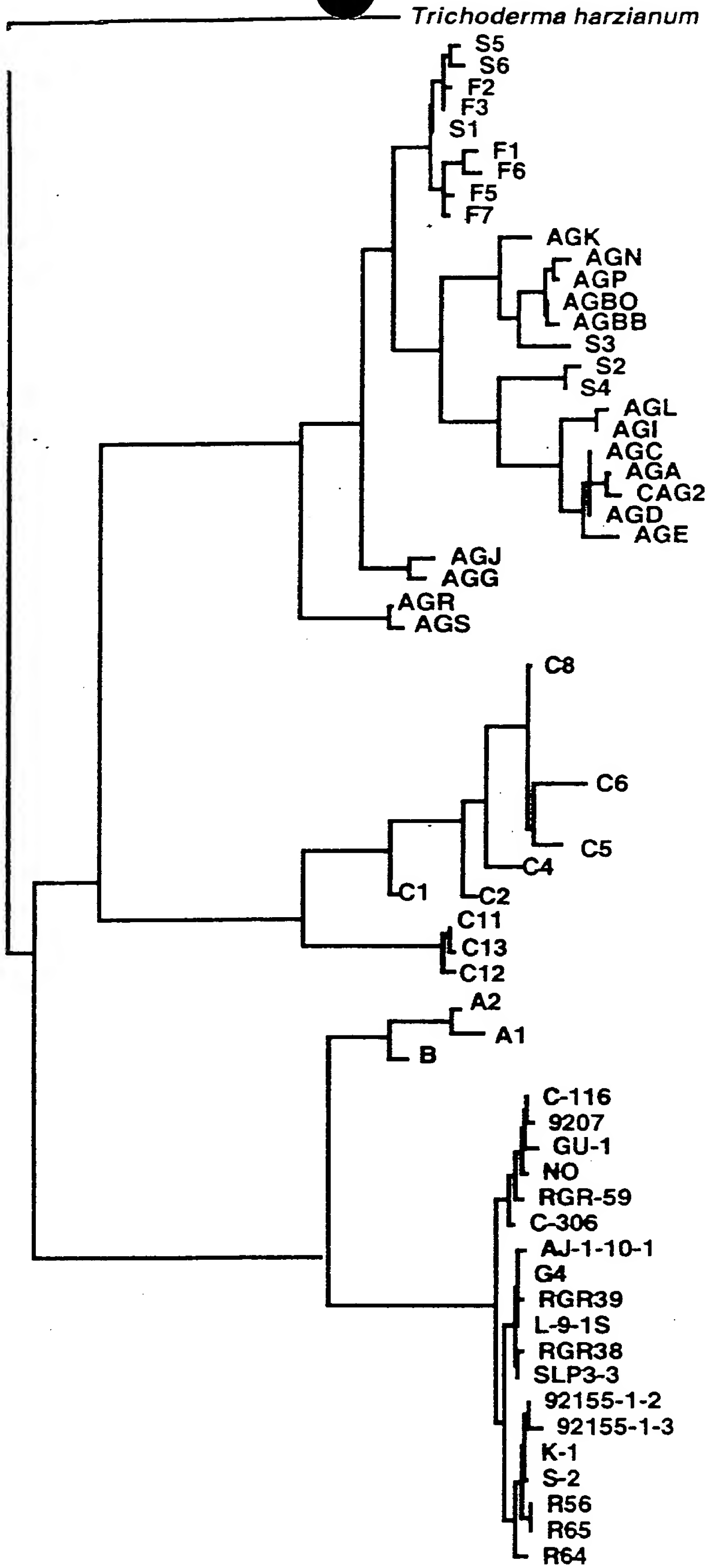
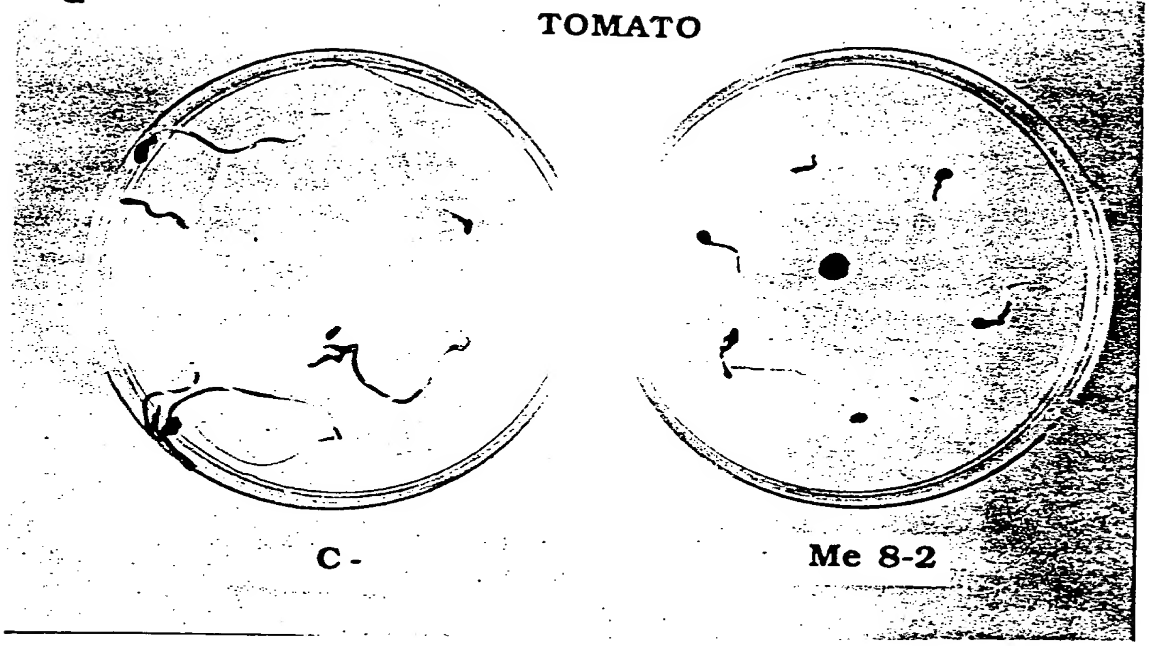

Trichoderma harzianum
S5
S6
F2
F3
S1
F1
F6
F5
F7
AGK
AGN
AGP
AGBO
AGBB
S3
S2
S4
AGL
AGI
AGC
AGA
CAG2
AGD
AGE
AGJ
AGG
AGR
AGS
C8
C6
C5
C4
C1
C2
C11
C13
C12
A2
A1
B
C-116
9207
GU-1
NO
RGR-59
C-306
AJ-1-10-1
G4
RGR39
L-9-1S
RGR38
SLP3-3
92155-1-2
92155-1-3
K-1
S-2
R56
R65
R64

Figure 2

Figure 3

a



b

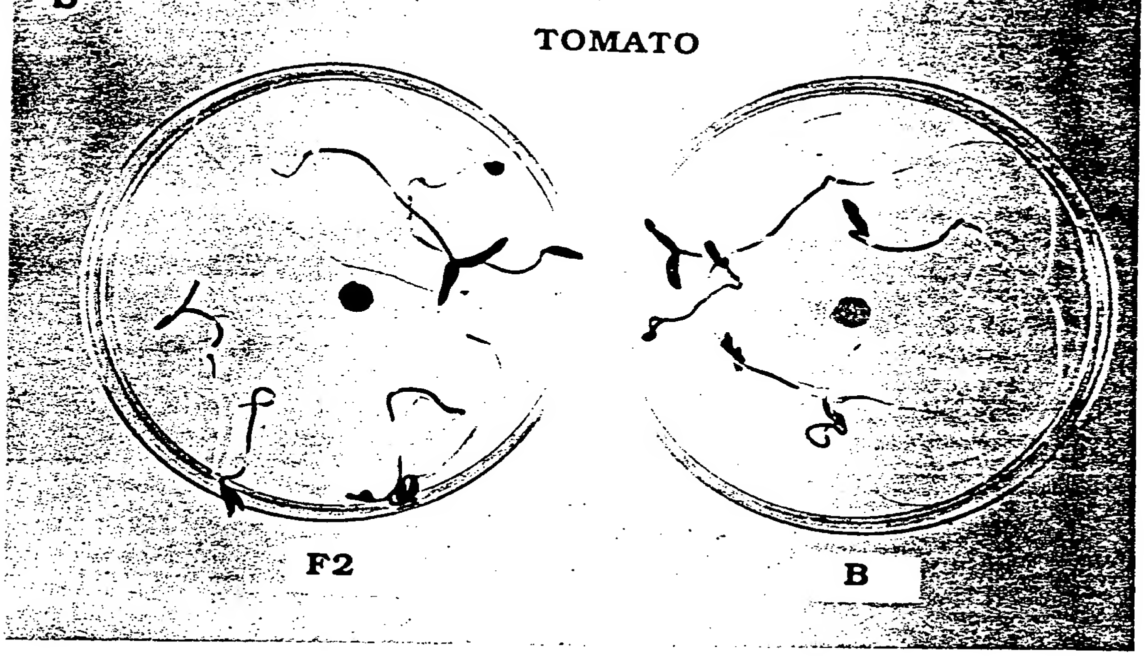
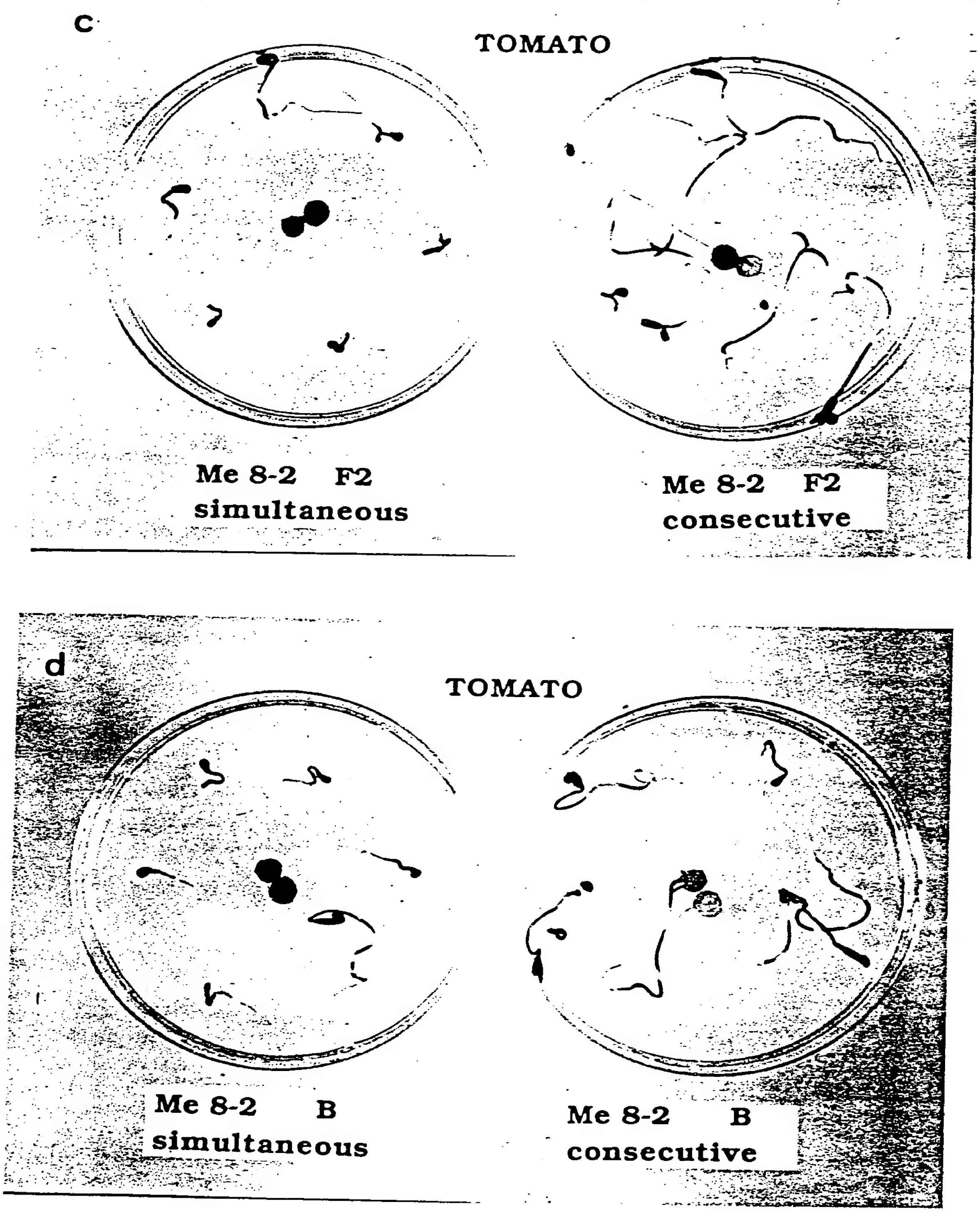


Figure 3


c
TOMATO
Me 8-2 F2
simultaneous
Me 8-2 F2
consecutive
d
TOMATO
Me 8-2 B
simultaneous
Me 8-2 B
consecutive

PCT/GB99/02406

marks + clerk

13/8/99

cp